



## **An atlas of O-linked glycosylation on peptide hormones reveals diverse biological roles**

Madsen, Thomas D.; Hansen, Lasse H.; Hintze, John; Ye, Zilu; Jebari, Shifa; Andersen, Daniel B.; Joshi, Hiren J.; Ju, Tongzhong; Goetze, Jens P.; Martin, Cesar; Rosenkilde, Mette M.; Holst, Jens J.; Kuhre, Rune E.; Goth, Christoffer K.; Vakhrushev, Sergey Y.; Schjoldager, Katrine T.

*Published in:*  
Nature Communications

*DOI:*  
[10.1038/s41467-020-17473-1](https://doi.org/10.1038/s41467-020-17473-1)

*Publication date:*  
2020

*Document version*  
Publisher's PDF, also known as Version of record

*Document license:*  
[CC BY](#)












*Citation for published version (APA):*  
Madsen, T. D., Hansen, L. H., Hintze, J., Ye, Z., Jebari, S., Andersen, D. B., Joshi, H. J., Ju, T., Goetze, J. P., Martin, C., Rosenkilde, M. M., Holst, J. J., Kuhre, R. E., Goth, C. K., Vakhrushev, S. Y., & Schjoldager, K. T. (2020). An atlas of O-linked glycosylation on peptide hormones reveals diverse biological roles. *Nature Communications*, 11(1), [4033]. <https://doi.org/10.1038/s41467-020-17473-1>

## ARTICLE


<https://doi.org/10.1038/s41467-020-17473-1>

OPEN

# An atlas of O-linked glycosylation on peptide hormones reveals diverse biological roles

Thomas D. Madsen <sup>1</sup>, Lasse H. Hansen <sup>1,2</sup>, John Hintze <sup>1</sup>, Zilu Ye <sup>1</sup>, Shifa Jebari<sup>3</sup>, Daniel B. Andersen<sup>4,5</sup>, Hiren J. Joshi <sup>1</sup>, Tongzhong Ju<sup>6</sup>, Jens P. Goetze<sup>2,4</sup>, Cesar Martin <sup>6</sup>, Mette M. Rosenkilde <sup>4</sup>, Jens J. Holst <sup>4,5</sup>, Rune E. Kuhre<sup>4,5</sup>, Christoffer K. Goth <sup>1</sup>, Sergey Y. Vakhrushev <sup>1</sup> & Katrine T. Schjoldager <sup>1</sup>✉

Peptide hormones and neuropeptides encompass a large class of bioactive peptides that regulate physiological processes like anxiety, blood glucose, appetite, inflammation and blood pressure. Here, we execute a focused discovery strategy to provide an extensive map of O-glycans on peptide hormones. We find that almost one third of the 279 classified peptide hormones carry O-glycans. Many of the identified O-glycosites are conserved and are predicted to serve roles in proprotein processing, receptor interaction, biodistribution and biostability. We demonstrate that O-glycans positioned within the receptor binding motifs of members of the neuropeptide Y and glucagon families modulate receptor activation properties and substantially extend peptide half-lives. Our study highlights the importance of O-glycosylation in the biology of peptide hormones, and our map of O-glycosites in this large class of biomolecules serves as a discovery platform for an important class of molecules with potential opportunities for drug designs.

<sup>1</sup>Copenhagen Center for Glycomics, Department of Cellular and Molecular Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3 DK-2200 Copenhagen N, Denmark. <sup>2</sup>Department of Clinical Biochemistry, Rigshospitalet, University of Copenhagen, Blegdamsvej 9 DK-2100 Copenhagen O, Denmark. <sup>3</sup>Biofisika Institute (UPV/EHU, CSIC), Departamento de Bioquímica, Universidad del País Vasco, Bilbao 48080, Spain. <sup>4</sup>Department of Biomedical Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3 DK-2200 Copenhagen N, Denmark. <sup>5</sup>Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Blegdamsvej 3 DK-2200 Copenhagen N, Denmark. <sup>6</sup>Office of Biotechnology Products, Center for Drug Evaluation and Research, U.S. Food and Drug Administration, Silver Spring, MD 20993, USA. ✉email: [Schjoldager@sund.ku.dk](mailto:Schjoldager@sund.ku.dk)

**P**eptide hormones, neuropeptides, and other biologically active peptides represent one of the largest and most promising classes of drug candidates. A wide range of synthetic analogs are emerging as major drug candidates for treatment of neurological and metabolic disorders, mediating their physiological effects by acting as either agonists or as antagonists on their cognate receptors<sup>1</sup>. While peptide hormone-based drugs benefit from being selective for their cognate receptors, the usage of native peptides as therapeutics is challenging due to the inherent instability giving rise to short half-lives of the biologically active peptide form(s) on a scale of a few minutes<sup>2</sup>. Thus, considerable efforts are being devoted to discovery and development of peptide-based analogs to improve stability and circulation time, as well as safety, immune tolerance, efficacy, and biodistribution. Different strategies include use of unnatural amino acids, modifications such as lipidation, PEGylation, conjugation to antibodies, and stapled peptides<sup>3</sup>. A prerequisite for rational design and development of improved peptide hormone drugs is knowledge of the molecular composition and structure of the natural bioactive peptides, and their mode of interaction with receptors.

Peptide hormones are synthesized as precursor proteins that may undergo regulated proteolytic processing during their transport through the secretory pathway to become mature hormones (involving enzymes such as carboxypeptidases<sup>4</sup>, pro-protein convertases (PCs)<sup>5</sup>, and C-terminal  $\alpha$ -amidation<sup>6</sup>). In addition, a number of other posttranslational modifications (PTMs) have been identified, including tyrosine sulfation<sup>7</sup>, N-terminal acetylation<sup>8</sup>, phosphorylation<sup>9</sup>, and glycosylation<sup>10</sup>. Unlike constitutively secreted proteins, most peptide hormones are stored within the cell. Upon secretion, most peptide hormones are prone to specific and rapid proteolytic degradation resulting in short circulatory half-lives<sup>11–13</sup>. As a result, the plasma concentration of most peptide hormones is in the lower picomolar range, which for technical reasons has challenged the peptidomic insight into the molecular composition of the mature forms of these hormones and their PTMs. Recent sensitive mass spectrometry (MS)-based studies have identified  $\alpha$ -amidation, acetylation, and phosphorylation on select peptide hormones<sup>9,14</sup>. However, given that the vast majority of proteins trafficking the secretory pathway undergo O-glycosylation, it is surprising that only very few O-glycosylated peptide hormones have been identified to date<sup>10,15–19</sup>.

Mucin-type O-glycosylation (hereafter O-glycosylation) is a highly abundant PTM estimated to affect 80% of secreted proteins in a cell<sup>15</sup>. The biosynthesis is initiated by a large family of 20 partly redundant polypeptide GalNAc-transferases (GalNAc-T) isoenzymes, and recently we and others have demonstrated that O-glycosylation plays wide roles in co-regulation of fundamental processes, such as PC processing, ectodomain shedding, and receptor modulation, to fine-tune and diversify biological functions of proteins<sup>20–22</sup>. Most recently, we identified O-glycosites on the atrial natriuretic peptide (ANP) hormone, and found that O-glycans in the receptor-binding region alter the mode of action of ANP in rats, providing improved therapeutic performance<sup>10</sup>. Prompted by this, we hypothesized that O-glycosylation could be a more prevalent PTM of peptide hormones than recognized to date, and that a focused discovery strategy would reveal this.

Here, we present the design and application of a comprehensive discovery program focused at identifying O-glycans on the large class of low molecular weight peptide hormones. We use different enrichment strategies and targeted MS-based analysis of multiple biosources to produce an atlas of O-glycosites on peptide hormones. We identify wide occurrence of O-glycans on select subfamilies of peptide hormones with multiple O-glycosites in the receptor-binding domains conserved among peptide sequences

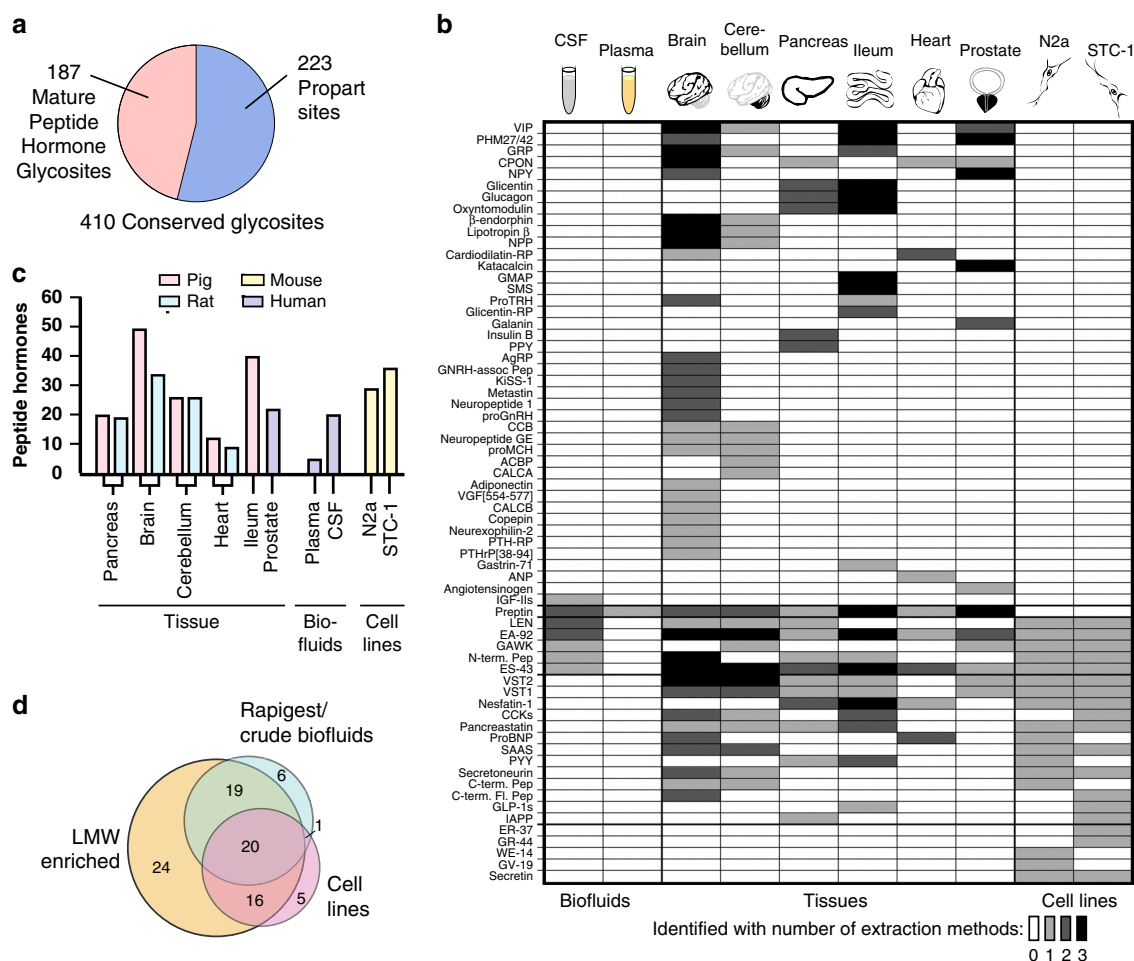
within subfamilies, as well as through evolution. We demonstrate that O-glycans in receptor-binding regions serve to modulate both peptide hormone stability and receptor signaling.

## Results

**Workflow for discovery of glycans on peptide hormones.** We designed a proteomics-based workflow targeting low molecular weight glycoproteins using low molecular weight enrichment (LMWE) and lectin weak affinity chromatography (LWAC) enrichments<sup>15,23</sup>, and applied a comprehensive database of neuropeptides and peptide hormones (NeuroPep; <http://isyslab.info/NeuroPep/>)<sup>24</sup> for informatics. The strategy (presented in Supplementary Fig. 1a) was applied to both cell lines (N2A and STC-1) and organs known to express and/or secrete high levels of diverse peptide hormones, including porcine and rat brain, cerebellum, heart, and pancreas as well as porcine intestinal ileum, and human plasma, cerebrospinal fluid (CSF), and prostate cancer biopsies (Fig. 1).

Across all biosources, we identified ~97,000 peptide-spectrum matches containing glycan information (Supplementary Fig. 1a). In total, 445 O-glycosites were found to be located on distinct prohormone proteins<sup>24</sup>. Of these, 36 sites were identified on human prohormone sequences and 409 glycosites were identified on nonhuman (porcine or rodent) prohormones. The latter were subjected to prediction of conservation in human neuropeptide orthologs by sequence alignment analysis and revealed that 374 out of 409 glycosites were predicted to be conserved in humans, based on conservative preservation of Ser/Thr residues within ( $\pm 5$ ) amino acid residues<sup>25</sup> (Supplementary Fig. 1b). In total, 410 (36 + 374) conserved O-glycosylation sites were included for further analysis. Surveying the distribution of the identified human and conserved O-glycosites across each peptide hormone precursor revealed that about half (223) of the 410 glycosites were located in the proprotein part, and the other half (187) were located on the mature peptide hormones (Fig. 1a, Table 1). These O-glycosites were distributed on 91 mature peptide hormone sequences out of the total 279 peptide hormones annotated in the comprehensive NeuroPep database<sup>24</sup> (including 14 peptide hormones for which we only observed ambiguously assignable glycosites; Fig. 1b, Table 1, Supplementary Fig. 1b, Supplementary Data 1–4). The identification correlated well with the known tissue expression of specific peptide hormones (e.g., insulin and pancreatic polypeptide (PPY) in pancreas). We identified fewest peptide hormones in plasma, and most in brain and intestinal tissues among the different biosources analyzed (Fig. 1c). In general, the LMWE resulted in more identifications of peptide hormone O-glycosites compared to the traditional analysis of total tissue/cell extracts (Rapigest) and biofluids, and in total contributed with >25% (24) additional peptide hormones with O-glycosylation (Fig. 1d). Each LMWE resulted in the identification of a similar number of glycosylated peptide hormones per biosource analyzed (Supplementary Data 2) where subtle qualitative differences were noted. For example, glycosylated insulin-like growth factor II (IGFII) was only identified in CSF after EtOH extraction, glycosylated glucagon-like peptide 1 (GLP-1) was found in ileum only after acidic extraction, and glycosylated CCB peptide was observed only after acetone extraction in both whole brain and cerebellum.

**An atlas of peptide hormone O-glycosylation.** Human peptide hormones are classified into 46 distinct families of which O-glycosylated members were identified in 28 of the families (Fig. 2, Supplementary Data 1–3). In addition, the algorithm NetOGlyc 4.0 (ref. <sup>15</sup>; <http://www.cbs.dtu.dk/services/NetOGlyc/>) predicted that the majority of peptide hormone families contain



**Fig. 1 Summary of the identification of glycosylated peptide hormones using LC-MS/MS. a** Number of glycosites identified on mature and propeptide hormone proteins. **b** Heatmap illustrating the identification of glycosites on mature peptide hormones across different biofluids, tissues, and cell lines. Greyscale intensity illustrates the number of extraction methods that have resulted in the identification of a given glycopeptide. See Supplementary Data 4 for examples of extracted spectra of glycopeptides from each peptide hormone. **c** Bar graph of the total number of mature glycopeptide hormones identified in each tissue or cell line (all enzymatic digestion procedures combined). **d** Venn diagram illustrating the distribution of identified glycosylated peptide hormones across biosources from the different extraction methods (total protein extractions (Rapigest extractions and crude biofluids), low molecular weight extractions (LMWE: H<sub>2</sub>O, acetone, acetone/AcOH, AcOH, and EtOH/HCl) and cell lines (supernatant and total cell lysate)). Please see Supplementary Data 1 for peptide hormone abbreviations.

glycosylated members (41 out of 46). Among the glycosylated peptide hormone families, well-characterized peptides like GLP-1, cholecystokinin (CCK), neuropeptide Y (NPY), galanin (GAL), and secretin (SCT) were found glycosylated.

We analyzed glycosite locations relative to protein topology and known functional features and found that the majority of sites were distributed among regions of the peptide hormones typically referred to as the address and message regions<sup>26</sup>. Approximately 10% of the glycosites were located  $\pm 5$  amino acids from a PC processing site (Fig. 3a). For example, in proopiomelanocortin (POMC; Supplementary Fig. 2a), we confirmed the Thr71 glycosite adjacent to the <sup>71</sup>TENPRK↓Y<sup>77</sup> processing site, which inhibits activation of gamma-melanocyte stimulating hormone<sup>27</sup>. In proNPY (Supplementary Fig. 2b), we identified Ser68 and Ser69 glycosites immediately adjacent to the PC1/3 activation (<sup>63</sup>RYGKR↓SS<sup>69</sup>) site in NPY<sup>28</sup>. In somatostatin (Supplementary Fig. 2c), an ambiguous glycosite (Ser89 or Ser92) was found just C-terminal to the furin/PACE4 activating site <sup>85</sup>ELQR↓SANS<sup>92</sup><sup>29</sup>.

We explored the address and message glycosites relative to characterized structural and/or functional receptor-binding motifs. Interestingly, in five different peptide hormone families

a majority of the glycosites were located in the highly conserved functional domains involved in receptor interaction (Fig. 3b). A detailed description of all five families is included in Supplementary Fig. 3.

### O-glycans on peptide hormones modulate receptor activation.

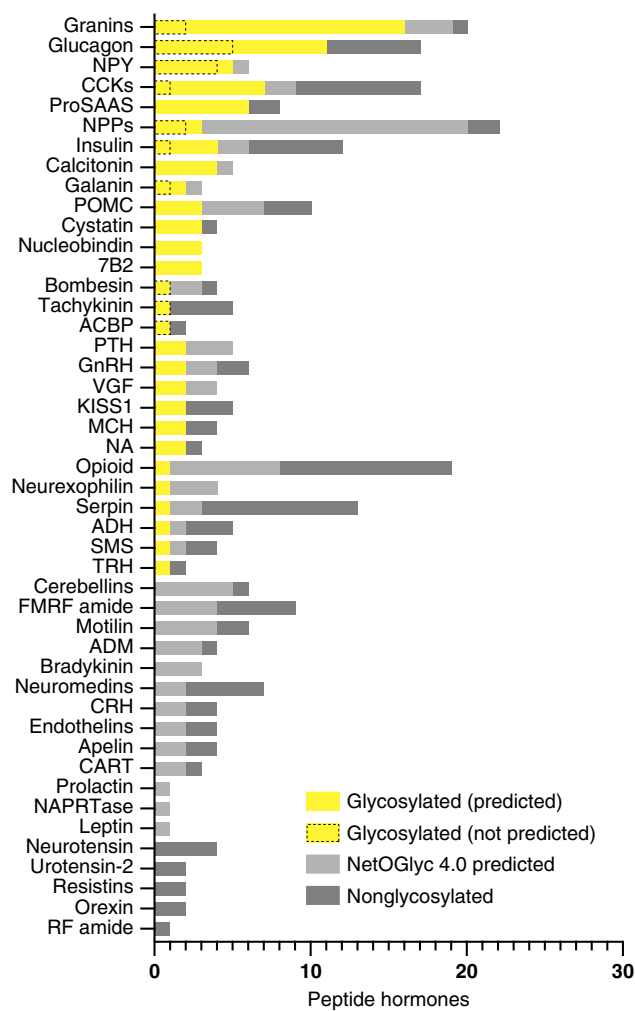
The glucagon (GCG) family members share a highly conserved O-glycan at Thr7, while the NPY family members share a conserved O-glycan at Thr32 (Fig. 4a, b). To explore the potential functional impact of site-specific O-glycosylation on mature peptide hormones, we chose to study members of these families (vasoactive intestinal polypeptide (VIP), GCG, GLP-1, NPY, and peptide YY (PYY)). We chemoenzymatically synthesized the candidate hormones with the three most common O-GalNAc-type structures, including Tn (GalNAc1-O-Ser/Thr), T (Gal $\beta$ 1-3GalNAc1-O-Ser/Thr), and sialylated ST (NeuAc $\alpha$ 2-3Gal $\beta$ 1-3GalNAc1-O-Ser/Thr; Fig. 4c), to generate NPY<sup>Thr32</sup>, PYY<sup>Thr32</sup>, VIP<sup>Thr7</sup>, GCG<sup>Thr7</sup>, and GLP-1<sup>Thr7</sup>. These peptide glycovariants were used for classical dose-response studies using COS-7 cells that transiently expressed relevant cognate receptors. All receptors were activated by the relevant non-modified peptide ligands

Table 1 Atlas of peptide hormone glycosylation.																
Family name	Peptide hormones	Brain pig	Brain rat	Cerebellum pig	Cerebellum rat	Heart pig	Heart rat	Pancreas pig	Pancreas rat	Ileum pig	Prostate human	CSF human	Plasma human	N2a mouse	STC-1 mouse	Phosphorylated <sup>a</sup>
7B2	C-terminal peptide	X	-	X	-	-	-	-	-	-	-	-	-	X	-	X
7B2	N-terminal peptide	X	X	-	-	-	-	-	-	X	-	X	-	X	X	X
7B2	Neuroendocrine protein 7B2	X	X	X	-	-	-	-	-	X	-	X	-	X	X	X
ACBP	Acyl-CoA-binding protein	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ADH	Copeptin	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
Bombesin	Gastrin-releasing peptide	X	X	-	X	-	-	-	-	X	-	-	-	-	-	-
Calcitonin	Calcitonin gene-related peptide 1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Calcitonin	Calcitonin gene-related peptide 2	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Calcitonin	Islet amyloid polypeptide	-	-	-	-	-	-	-	X	-	-	-	-	-	X	-
Calcitonin	Katacalcin	-	-	-	-	-	-	-	-	-	X	-	-	-	-	-
CKCs	Cholecystokinin-25/33/39/58/(desnopeptide)	X	-	X	-	-	-	-	-	X	-	-	-	-	X	-
CKCs	Gastrin-71	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grans	CCB peptide	X	-	X	-	-	-	-	-	-	-	-	-	-	-	-
Grans	EA-92	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	ES-43	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Grans	GAWK peptide	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grans	GR-44/ER-37	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grans	GV-19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grans	Pancreastatin	-	X	-	X	-	-	-	-	X	-	-	-	-	-	-
Grans	Chromogranin-A	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Grans	Secretogranin-1	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	Secretogranin-2	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	Secretogranin-3	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	Secretoneurin	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	Vasostatin-1	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	Vasostatin-2	X	X	X	X	X	-	X	-	X	X	X	X	X	X	X
Grans	WE-14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Grans	Kininogen-1	-	X	-	X	-	X	-	X	X	X	X	X	-	-	-
Cystatin	Kininogen-1 heavy chain	-	-	-	-	-	-	-	-	X	X	X	X	-	-	-
Cystatin	Kininogen-1 light chain	-	X	-	X	-	X	-	X	X	X	X	X	-	-	-
Galanin	Galanin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Galanin	Galanin message-associated peptide	-	-	-	-	-	-	-	-	X	-	-	-	-	-	X
Glucagon	Glucagon	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-
Glucagon	Glucagon-related polypeptide	-	-	-	-	-	-	-	-	X	-	-	-	-	-	-
Glucagon	Glucagon-like peptide 1's	-	-	-	-	-	-	-	-	X	-	-	-	-	X	-
Glucagon	Glucagon/glicentin	-	-	-	-	-	-	X	X	X	X	-	-	-	-	-
Glucagon	Intestinal peptide PHM-27/PHV-42	X	X	-	-	-	-	-	-	X	X	-	-	-	-	-
Glucagon	Oxyntomodulin	-	-	-	-	-	-	X	X	X	-	-	-	-	-	-
Glucagon	Secretin	-	-	-	-	-	-	-	-	-	-	-	-	X	-	-
Glucagon	Vasoactive intestinal peptide	X	X	X	-	-	-	-	-	X	X	-	-	-	-	-
GnRH	GnRH-associated peptide 1	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GnRH	Progonadoliberin-1	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Insulin	Insulin B chain	-	-	-	-	-	-	X	X	-	-	-	-	-	-	-
Insulin	Insulin-like growth factor II/Ala-25 Del	-	-	-	-	-	-	-	-	-	-	X	X	-	-	-
Insulin	Preptin	X	-	X	-	-	-	X	-	X	-	-	-	-	-	-
Insulin	Metastasis-suppressor KISS-1	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KISS1	Metastin	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
KISS1	Neuropeptide-glycine-glutamic acid/pro-MCH	-	X	-	X	-	-	-	-	-	-	-	-	-	-	-
MCH	Adiponectin	-	X	-	-	-	-	-	-	-	-	-	-	-	-	-
NA	Agouti-related protein	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NA	Agouti-related protein	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Natriuretic peptide	Atrial natriuretic factor	-	-	-	-	X	X	-	-	-	-	-	-	-	-	-
Natriuretic peptide	Cardiolipin-related peptide	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Natriuretic peptide	Natriuretic peptides B	X	-	X	X	-	-	-	-	-	-	-	-	-	-	-
Neurexophilin	Neurexophilin-2	X	-	-	-	-	-	-	-	-	-	-	-	-	-	-
NPY	C-flanking peptide of NPY	X	X	-	-	-	-	-	X	-	X	-	-	-	-	X
NPY	Neuropeptide Y	X	-	-	-	-	-	-	-	-	X	-	-	-	-	X
NPY	Pancreatic hormone (PPY)	-	-	-	-	-	-	X	X	-	-	-	-	-	-	-
NPY	Peptide YY/(3-36)	-	-	-	-	-	-	X	X	-	-	-	-	-	-	-
Nucleobindin	Nesfatin-1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Nucleobindin	Nucleobindin-1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Nucleobindin	Nucleobindin-2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Nucleobindin	Neuropeptide 1	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-
Opioid	Parathyroid hormone	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-
ProSAA5	Beta-endorphin/lipotropin beta/NPP	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-
ProSAA5	Big LEN/little LEN/ big PEN-LEN	X	X	-	-	-	-	-	-	-	-	-	-	-	-	-
ProSAA5	Big SAA5/little SAA5	X	-	X	-	-	-	-	-	-	-	-	-	-	-	-

Table 1 (continued)													
Family name	Peptide hormones	Brain pig	Brain rat	Cerebellum pig	Cerebellum rat	Heart pig	Heart rat	Pancreas pig	Pancreas rat	Ileum pig	Prostate human	CSF human	Plasma human
ProSAAS	proSAAS	X	X	X	X	-	-	-	X	-	-	-	-
Serpin	Angiotensinogen	-	-	-	-	-	-	-	-	-	-	-	-
Somatostatin	Somatostatin-28	-	-	-	-	-	-	-	-	-	-	-	-
Tachykinin	C-terminal-flanking peptide	-	-	-	-	-	-	-	-	-	-	-	-
TRH	Pro-thyrotropin-releasing hormone	X	X	-	-	-	-	-	-	-	-	-	-
VGF	Antimicrobial peptide VGF[554-577]	X	X	-	-	-	-	-	-	-	-	-	-
VGF	Neurosecretory protein VGF	X	X	X	X	-	-	-	-	-	-	-	-

X indicates the identification of peptide hormone glycosylation in a given biosource. - indicates no identification in a given biosource (See Supplementary Data 1 for full atlas of glycosylated peptide hormones). Please see Supplementary Data 1 for peptide hormone and family abbreviations.

aPeptide hormone phosphorylation sites previously identified<sup>9,14</sup> (Supplementary Data 5).

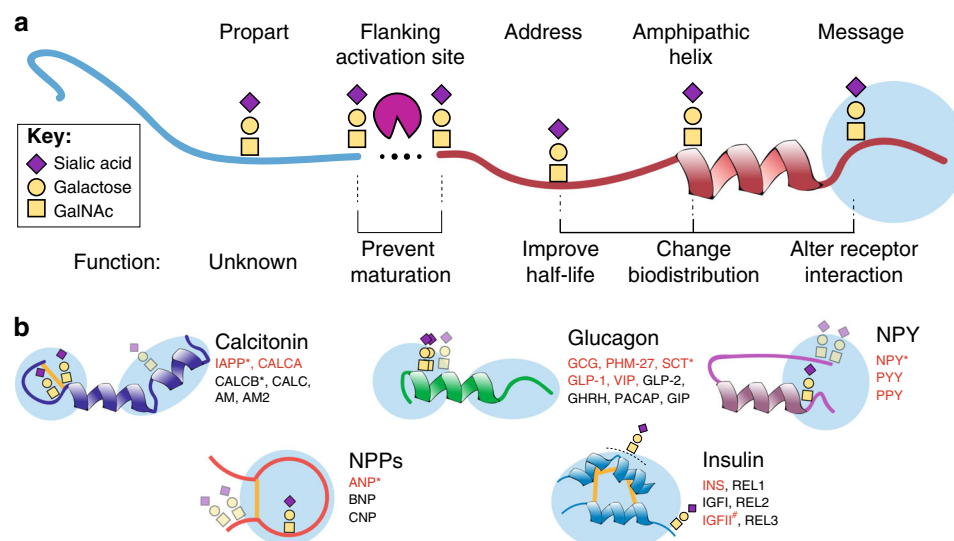


**Fig. 2 Glycosylation of peptide hormones is a widespread phenomenon.** Bar graph illustrating that 28 out of the 46 human families annotated in the NeuroPep<sup>24</sup> database contain at least one glycosylated member (yellow). The glycosylation prediction algorithm, NetOGlyc 4.0, predicts that almost all (41 out of 46) families contain glycopeptide hormones (light gray) and five families only have members without glycosylation (dark gray). The number of identified peptide hormones that NetOGlyc 4.0 does not predict are marked by a dotted line. Please see Supplementary Data 1 for peptide hormone family abbreviations.

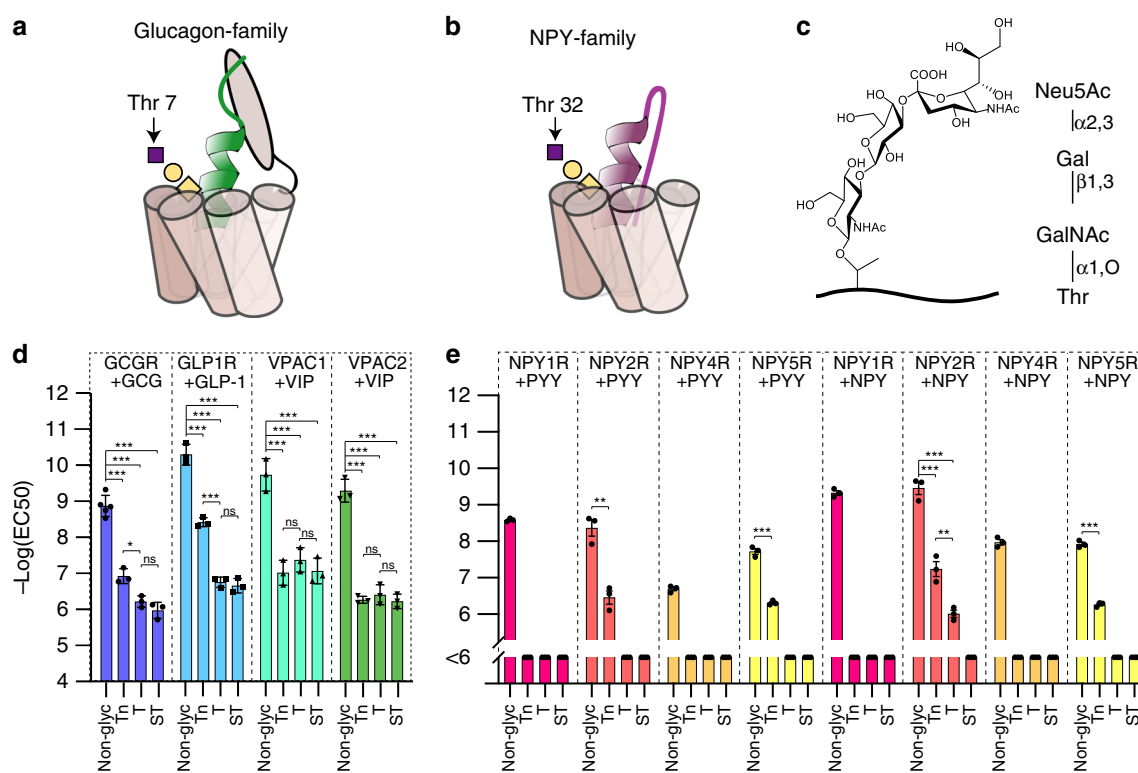
producing half the maximal effect ( $EC_{50}$ ) at doses in accordance with the literature<sup>30–33</sup>. A summary of all dose–response curves is presented in Supplementary Fig. 4, and exact  $EC_{50}$  values can be found in Supplementary Table 1.

In the GCG family, non-glycosylated VIP exhibited a potency of 0.2 nM and 0.4 nM for VIP receptor 1 and 2 (VPAC1 and 2), respectively. In striking contrast, all three glycoforms of VIP<sup>Thr7</sup> showed a strongly reduced potency on both VPAC1 and VPAC2 (Fig. 4d). GLP-1 signals exclusively through the GLP-1 receptor (GLP-1R), and in our assay non-glycosylated GLP-1 showed a potency of 0.05 nM for GLP-1R, while GLP-1<sup>Thr7</sup> produced a less potent yet full maximal response. Activation of the glucagon receptor (GCGR) by GCG occurred with an  $EC_{50}$  of 1.3 nM, while GCG<sup>Thr7</sup> exhibited decreased potency with 100–1000-fold increased  $EC_{50}$  dependent on the structure of the installed O-glycan (Fig. 4d, Supplementary Figs. 4 and 5a). The NPY family peptide hormones activate the same members of the Gi-coupled NPY receptor family (NPY1R, NPY2R, NPY4R, and NPY5R; Fig. 4e, Supplementary Fig. 4). In agreement with past studies,

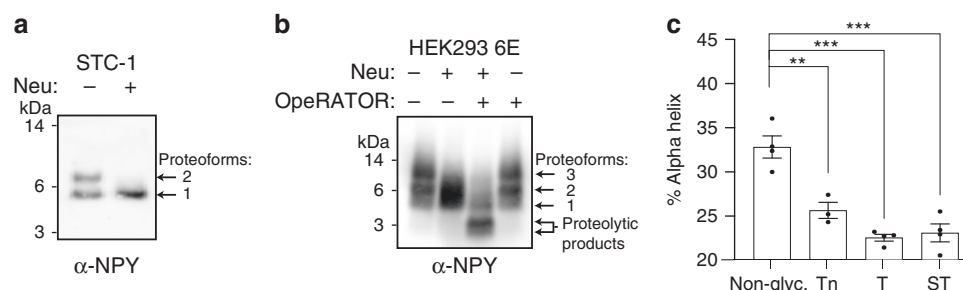




**Fig. 3 Topology of O-glycosites in peptide hormones and their predicted functions.** **a** Representation of a generic peptide hormone with indications of potential functional impact of glycosylation sites. **b** Illustrations of glycopeptide hormone families with message-associated O-glycosites conserved or semi-conserved among the paralogous family members. The message parts of the peptide hormones are indicated by light blue circles. Glycans have been identified in conserved positions in family members in red. Faded glycosites are not shared within paralogous family members. (\*) Denotes peptide hormones containing non-conserved glycosites. (#) The site over the dotted line in the insulin family is covered by ambiguously assigned glycosite(s) in the C-terminal part of IGFII. Full sequence alignments of the five presented glycopeptide hormone families are shown in Supplementary Fig. 3. Please see Supplementary Data 1 for peptide hormone abbreviations.



**Fig. 4 Functional analysis of O-glycans within the receptor-binding region of peptide hormones.** **a** Schematic depiction of the Thr<sup>7</sup> O-glycan positioned within the receptor-binding domain of the glucagon hormone family (see Supplementary Fig. 7e). **b** Schematic depiction of the Thr<sup>32</sup> O-glycan positioned within the receptor-binding domain of the NPY hormone family<sup>38</sup>. **c** Structural representation of the sialyl-T O-glycan. **d** Potencies ( $\text{Log}_{10}(\text{EC}_{50})$ ) of glycosylated GCG, VIP, and GLP-1 determined in cell-based receptor activation assays ( $n = 3$  independent experiments in duplicate determinations for all groups, except non-glyc. GCG where  $n = 5$ ). **e** Potencies of NPY and PYY in their Thr<sup>32</sup> glycoforms determined and shown as in **d** ( $n = 3$  for all groups). All error bars represent mean  $\pm$  S.E.M. Two-tailed one-way ANOVA (Tukey's post hoc test) was performed in **d** and **e** except for comparisons within NPY2R + PYY-; NPY5R + PYY-, and NPY5R + NPY- assays in **e** where two-tailed Student's  $t$ -test was performed. ns not significant; \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$  (all significant  $p$ -values are  $< 0.0001$  except **d** GCGR-Tn/GCGR-T  $p = 0.0149$ ; **e** NPY2R-PYY-Non-glyc/NPY2R-PYY-Tn  $p = 0.0027$ ; NPY2R-NPY/NPY2R-NPY-Tn  $p = 0.0002$ ; NPY2R-NPY-Tn/NPY2R-NPY-T  $p = 0.0046$ ). Source data for **d** and **e** are provided as a Source data file.



**Fig. 5 Molecular and structural analysis of NPY and its precursor.** **a** Western blot of conditioned media from STC-1 cells stably expressing proNPY using an antibody recognizing pro- and mature NPY. The conditioned media from proNPY-transfected STC-1 cells was concentrated using TCA precipitation, and analyzed with and without neuraminidase treatment (Neu). **b** Western blot of conditioned media from proNPY-transfected HEK293-6E cells were analyzed with and without neuraminidase treatment (Neu), OPERATOR treatment, or both in combination. Both western blot experiments in **a** and **b** were repeated three times with similar results. **c** Alpha-helical content of 20  $\mu$ M NPY or its glycoforms determined by CD spectroscopy (pH 7 at 25 °C) ( $n = 4$  independent measurements for all groups except Tn where  $n = 3$ ). All error bars represent mean  $\pm$  S.E.M. Two-tailed one-way ANOVA (Tukey's post hoc test) was performed in **c**. \*\*\* $p < 0.001$ , \*\* $p < 0.01$  (all  $p$ -values are  $< 0.0001$  except **c** Non-glyc/Tn where  $p = 0.0018$ ). Source data for **a–c** are provided as a Source data file.

non-glycosylated PYY and NPY activated receptors with potencies ( $EC_{50}$ ) between 2.6–204 nM and 0.4–10.7 nM, respectively<sup>34,35</sup>. One exception was activation of the Y5 receptor that showed an  $EC_{50}$  value of 204 nM compared to previously reported 1.2 nM<sup>34</sup>. Interestingly, among the four receptors tested, NPY2R and NPY5R partially tolerated NPY<sup>Thr32</sup>-Tn and PYY<sup>Thr32</sup>-Tn with  $EC_{50}$  values in the range of 58–550 nM (Fig. 4e), whereas NPYR1 and NPYR4 showed only slight activation at the highest dose of the glycosylated peptide hormones tested (Supplementary Figs. 4 and 5b).

In summary, all O-glycosylated peptide hormones elicited a dampened agonist response with increased  $EC_{50}$  values correlating positively with the size of the O-glycan. The presence of sialic acid with its strong negative charge, however, did not substantially affect the potency of the members of the GCG family when comparing to the other tested variants. In general, O-glycans located in the receptor-binding domain of the NPY family members had a bigger impact on receptor activation than glycans in the receptor-binding domain of GCG family members (Fig. 4d, e). Further investigation is required to explain this phenomenon, but it could be related to the large extracellular N-terminal domain of the receptors to the GCG family members<sup>36</sup> (like other class B G-protein-coupled receptors), which may allow for binding of the C-terminally non-glycosylated part (Fig. 4a). In contrast, NPY receptors (which belong to A-class G-protein-coupled receptors) do not contain this large extracellular N-terminal domain<sup>37</sup> (Fig. 4b).

#### ProNPY is highly glycosylated in a neuroendocrine cell line.

Estimation of stoichiometry using LC-MS/MS is difficult with LWAC-enriched glycopeptide samples, because the non-glycosylated peptide fraction is depleted by the lectin enrichment. In addition to this, the determination of glycosylation stoichiometry in non-enriched samples is challenged by the general sample complexity and dynamic range<sup>38</sup>. Studying the low abundant peptide hormones that are often not detectable in proteomics studies further adds to the difficulties in detection of glycopeptides let alone identification of glycoforms and stoichiometry of glycosylation. Thus, to address stoichiometry with one peptide hormone, we stably expressed full-length preproNPY in the neuroendocrine STC-1 cell line and used traditional SDS-PAGE molecular weight shift assays to evaluate the degree of O-glycosylation. Direct SDS-PAGE western blotting of trichloroacetic acid concentrated conditioned media showed that secreted proNPY migrated as two distinct and equally intense

bands (proteforms 1 and 2, Fig. 5a, lane 1). Treatment with neuraminidase to remove sialic acids on glycans collapsed the two immunoreactive bands of proNPY into one apparently co-migrating band, which demonstrated that the upper band of proNPY contains sialylated glycans (lane 2). To confirm the presence of GalNAc-type O-glycans, we expressed proNPY in a high producer cell line, HEK293-6E, that secreted proNPY in amounts directly detectable by western blotting without concentrating the sample further. HEK293-6E-produced proNPY migrated as three bands; two upper strong (proteform 2 and 3) and one lower weak (proteform 1) band (Fig. 5b, lane 1). As with STC-1, neuraminidase demonstrated that a considerable amount of proNPY contains sialylated glycans (lane 2). For further confirmation, we used a recently characterized O-glycoprotease from *Akkermansia muciniphila*, OPERATOR, which specifically cleaves O-glycopeptides immediately N-terminal to an O-glycan, but is blocked by the presence of sialic acid<sup>39</sup>. Treatment with OPERATOR only in combination with neuraminidase resulted in the loss of the intensity of the upper band concomitantly with the appearance of a band migrating approximately as 4–5 kDa (lane 3), confirming that a majority of proNPY contains one or more O-glycans with sialic acid. In our atlas, we identified multiple glycosylation sites on proNPY, and based on this data it was not possible to determine the specific OPERATOR cleavage site(s) (Supplementary Fig. 6). However, the data confirmed that two mammalian cell lines, STC-1 and HEK293-6E, produce proNPY with at least 50% sialylated GalNAc-type O-glycans.

#### Glycans in mature NPY unfolds its amphipathic alpha-helix.

GalNAc-type O-glycans are most often found in unstructured regions of glycoproteins<sup>15,40</sup>. Surprisingly, the conserved glycans in all the members of the GCG, NPY, and calcitonin families are positioned at the C- or N-terminus of amphipathic alpha-helices (Supplementary Fig. 7a–c). The alpha-helix is important for receptor recognition and activation<sup>41</sup>, and specifically for NPY, the alpha-helix is suggested to play a crucial role in prompting a two-step receptor-binding mechanism. Here, NPY first binds to the lipid membrane to increase its effective concentration<sup>41,42</sup>, followed by a lateral diffusion on the membrane to the receptor, where NPY binds its receptors and activates intracellular signaling cascades. To explore the potential structural impact of O-glycans on NPY, we measured mature NPY in solution without and with O-glycans at Thr32 using circular dichroism (CD) spectroscopy. NPY showed 33% alpha-helical content in line with



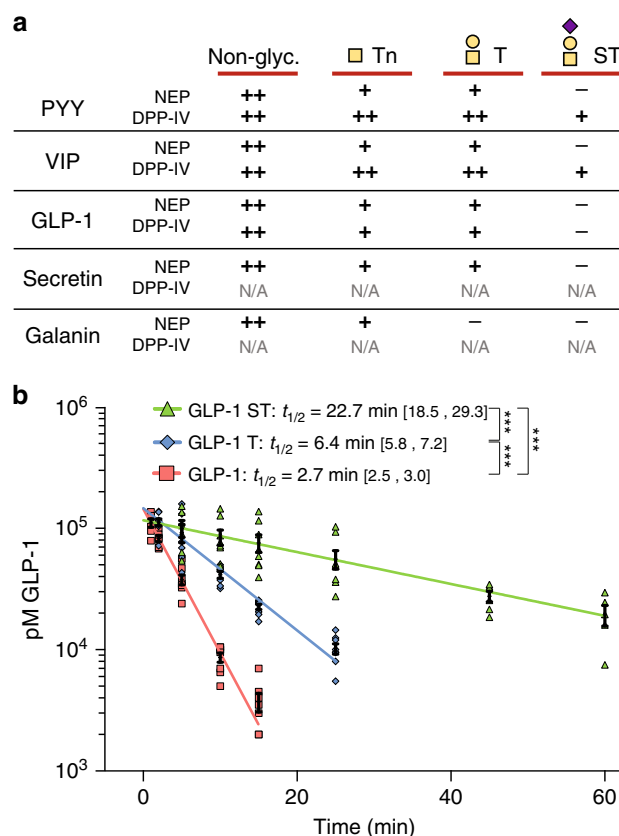
previous studies<sup>43</sup>, whereas NPY<sup>Thr32</sup> glycovariants with O-glycans in the alpha-helical interface decreased the alpha-helical content to 23–26%, seen as an increase in ellipticity at 222 nm (Fig. 5c, Supplementary Fig. 7d). Interestingly, where the Tn structure produced a major change, elongation to T and ST structures did not further change the secondary structure dramatically, which is in line with previous studies of mucin O-glycopeptides<sup>44</sup>.

**O-glycans modulate peptide hormone stability in vitro.** Most peptide hormones are rapidly degraded with half-lives of only a few minutes<sup>11,13</sup>. O-glycans in close proximity to proteolytic processing sites can modulate the rate of processing<sup>20</sup>. To study the effects of O-glycans on peptide hormones, we subjected selected glycosylated peptide hormones to in vitro degradation by neprilysin (NEP) or dipeptidyl peptidase IV (DPP-IV) enzymes, which are known to metabolize a large number of peptide hormones and other bioactive peptides in vivo, including PYY, VIP, GLP-1, SCT, and GAL<sup>45,46</sup>. Both DPP-IV and NEP degraded all tested non-glycosylated peptides fully or partially within 15–30 min, while the corresponding glycopeptides demonstrated increased stability for 60–120 min (Fig. 6a, Supplementary Fig. 8). In some cases, the intact, non-degraded glycopeptide hormones were still detectable even after overnight incubations. The protective effects seemed to be most profound with the larger sialyl-T glycoforms.

**GLP-1<sup>Thr7</sup> exhibits increased circulation time in vivo.** Intact GLP-1 (7–36 amide) is rapidly cleared from circulation with a half-life of ~2 min<sup>47</sup>. To evaluate the effect of O-glycosylation on GLP-1 clearance in vivo, we measured the half-life of GLP-1<sup>Thr7</sup> glycoforms after administration of an i.v. bolus (20 nmol/kg) of GLP-1 to male Wistar rats ( $n = 8$ ; Fig. 6b). In agreement with previous studies, GLP-1 and immunoreactive metabolites were rapidly cleared from circulation with a half-life of  $t_{1/2} = 2.7$  min when fitted to a one-phase decay model, while GLP-1<sup>Thr7</sup> glycoforms showed significantly slower elimination with half-lives of  $t_{1/2} = 6.4$  min (T-variant) and  $t_{1/2} = 22.7$  min (ST variant), respectively. As a result, non-glycosylated GLP-1 was undetectable after the 15 min timepoint from injection, while GLP-1<sup>Thr7</sup>-ST remained readily detectable in circulation even after 60 min.

## Discussion

Here, we developed a strategy to explore O-glycosylation on peptide hormones and demonstrate that almost 33% of all known human peptide hormones may carry O-glycans in four mammalian species investigated. We present a comprehensive atlas of O-glycosites on the large class of peptide hormones, which should find use in wider studies into specific functions of O-glycans on individual hormones. We highlight examples of biological function of site-specific O-glycans in select peptide hormones, demonstrating that O-glycans can modify receptor activation and prevent degradation and clearance. GalNAc-type O-glycosylation is an exceptional form of protein glycosylation in that it is regulated by a large family of up to 20 GalNAc-T isoforms, providing opportunities for a high degree of differential regulation of glycosites on specific proteins in cells<sup>48</sup>. The GalNAc-T isoforms are known to serve co-regulatory roles in fine-tuning functions of specific proteins with clear disease-causing consequences<sup>23,49–51</sup>, and we recently demonstrated that graded induction of individual GalNAc-T isoforms in cell models provides corresponding graded O-glycosylation of highly select glycosite substrates specific to the induced GalNAc-T<sup>52</sup>. In the present study, we did not address the biosynthesis of glycosylated peptide hormones and in particular the specific GalNAc-T



**Fig. 6** Stability of glycosylated peptide hormones in vitro and in vivo. **a**

Summary of in vitro degradation of peptide hormones using purified NEP or DPP-IV monitored by matrix assisted laser desorption/ionization-time of flight (MALDI-TOF) timecourse assay (See Supplementary Fig. 8 for spectra). Qualitative scoring ++: fully susceptible to degradation, +: reduced susceptibility, –: not susceptible to degradation. **b** Plasma concentrations of GLP-1 or glycoforms over time after intravenous bolus injection (20 nmol per kg) in anesthetized rats ( $n = 8$  rats per group). GLP-1 was measured by an assay targeting the amidated C-terminus of the molecule measuring total GLP-1 (intact 7–36 amide and metabolites hereof). Half-lives ( $T_{1/2}$ ) are indicated with 95% confidence intervals. For each glycoform, timepoints represented by  $n > 4$  measurements were included in the analysis ( $n_{\text{GLP-1}} = 35$ ,  $n_{\text{GLP-1 T}} = 38$ ,  $n_{\text{GLP-1 ST}} = 41$  measurements in total). Multiple linear regression was performed in **b**. ns not significant; \*\*\* $p < 0.001$  (all indicated  $p$ -values  $< 0.0001$ ). Source data for **b** is provided as a Source data file.

isoforms controlling glycosylation of the identified glycosites. This is a considerable challenge for future studies that should be aided by the presented atlas of glycosites.

Surprisingly, we identified >150 glycosites on mature peptide hormones with many located in the receptor-interacting regions. O-glycans are known to modulate receptor functions, as first described for O-fucosylation of the Notch receptor<sup>53</sup>, and more recently for GalNAc-type O-glycosylation of the low-density lipoprotein receptor-related receptors<sup>54–56</sup>. However, to our knowledge, receptors recognizing protein/peptide ligands have not yet been described to exhibit specific recognition of O-glycosylated ligands, apart from the large families of lectin receptors that primarily recognize the glycans<sup>57</sup>. In agreement with this, we found that O-glycans at specific well-conserved amino acids in VIP, GCG, GLP-1, NPY, PYY, and PPY resulted in lower receptor potency in vitro. For GCG and GLP-1, the size of the O-glycan correlated positively with the observed reduction in potency and for VIP all glycoforms diminished activation

similarly (Fig. 4). Among the NPY receptors tested, NPY2R and NPY5R were less affected by O-glycans on NPY and PYY, while O-glycans completely abolished NPY1R and NPY4R activation. Chemical modifications of amino acids in the receptor-binding domain of NPY are known to alter selectivity for the different receptors<sup>58</sup>, and our results showing that O-glycans in the ligand-binding region of NPY altered the receptor interaction are well in agreement with this. Although, with the lectin enrichment strategy, we gain the sensitivity needed for glycosite identification, we lose the opportunity for broad assessment of O-glycan stoichiometry. Stoichiometric analysis of O-glycans is one of the major challenges for the field; however, we did demonstrate high glycosylation occupancy for proNPY when expressed in two different cell lines. We expect that future studies will reveal how this reflects the O-glycan occupancy found in circulation or at the local site of secretion.

Interestingly, we observed a substantial overlap between reported phosphorylated peptides from two recent peptidomics studies<sup>9,14</sup>, and the identified O-glycosites on peptide hormones (24 out of 91 glycosylated peptide hormones contain both PTMs, see Table 1 and Supplementary Data 5). In several peptide hormones, we found closely positioned phosphosites (superscript P) and O-glycosites (superscript G), including NPY (NLIT<sup>P/G</sup>RQRY<sup>P</sup> and S<sup>G</sup>GPETLISDLLMRESPT<sup>G/P</sup>ENVPR<sup>T</sup>RLDPA), somatostatin (QRS<sup>P/G</sup>ANS<sup>P/G</sup>NPA), and vasostatin-2 (VEEPS<sup>P</sup>SGKDV<sup>M</sup>). This may suggest an interplay between the two PTMs much like the illustrative example of the fibroblast growth factor FGF23<sup>21,59</sup>. Here, O-glycosylation by the GalNAc-T3 isoform inhibits PC processing by furin and adjacent phosphorylation by the Fam20C kinase inhibits glycosylation, and augments processing and inactivation. Thus, glycosylation and phosphorylation may regulate the function of peptide hormones in concert with each other.

Peptide hormones are inherently unstable and circulate in minute amounts. Here, we demonstrated that O-glycans on SCT, VIP, GAL, GLP-1, and PYY protect the peptide hormones from proteolytic degradation in vitro. The most prominent effects were observed with the naturally occurring sialylated ST O-glycan structure, where, e.g., full-length PYY<sup>Thr32</sup> and GAL<sup>Ser23</sup> remained readily detectable after 24 h incubation with NEP. In addition, the in vivo circulatory half-life of GLP-1 was increased approximately tenfold compared to the non-glycosylated peptide. These results are well in line with the general concept that protein glycosylation increases stability by shielding the protein backbone from proteolytic cleavage, and also that sialic acid capping of glycans is needed to maintain plasma half-life by, e.g., reducing uptake by the Ashwell–Morell (asialoglycoprotein) receptor<sup>60,61</sup>. Although not addressed here, protein glycosylation can change the cellular uptake and, as shown for an enkephalin-derived glycopeptide, increase both transport across the blood–brain barrier and general bioavailability<sup>62</sup>. Whether naturally occurring O-glycans on peptide hormones affect biodistribution in vivo remains to be elucidated. Our data suggests that O-glycosylation of peptide hormones serves to fine-tune their bioactivities by decreasing receptor stimulation and increasing bioavailability, which may be advantageous for inducing lower and sustainable signals. The O-glycans may also be important for producing selective signals for subsets of related receptors, as we recently described in case of ANP<sup>10</sup>. The NPY family members are ubiquitously expressed in the body, and act as neurotransmitters to regulate a vast array of physiological processes via binding and signaling through the Gi-coupled NPY receptors (NPY1R, NPY2R, NPY4R, or NPY5R), and PTMs may play a role in orchestrating selectivity as observed for the N-terminal truncation by DPP-IV<sup>34</sup>. Another example is GLP-1, whose primary functions include increase of insulin secretion (i.e., to act as an incretin), inhibition of gastrointestinal

motility and physiological regulation of appetite, where PTMs may be involved in directing the response.

The great therapeutic potential of peptide hormones was realized almost a century ago and candidate drug designs continue to emerge<sup>2</sup>. Basing peptide therapeutics on naturally occurring biomolecules is a favorable approach since this often leads to lower toxicity and immunogenicity, as well as higher selectivity, and thus predictable physiological actions and behavior<sup>2</sup>. However, probing the function and efficacy of unmodified peptides has demonstrated low biostability and circulation time, and therefore low efficacy<sup>2,63</sup>. A number of studies have explored the chemical introduction of artificial glycans on peptide hormones<sup>64–67</sup>, and interestingly all studies find—very much in line with what we show here for the naturally occurring O-glycans—that the analogs chemically modified with artificial glycan structures are less or equally potent in in vitro receptor activation assays, yet more potent and stable in vivo.

Peptide hormones are frequently detected in clinical immunoassays and the presence of O-glycans on peptide hormones may mask antibody epitopes, which leads to underestimated concentrations as previously demonstrated for NT-proBNP<sup>68,69</sup>. Thus, the atlas may also serve as a guide in the strategic design of biomarker assays.

In summary, we provide a comprehensive atlas of O-glycosites on peptide hormones, and illustrate examples of functions of site-specific O-glycans for processing, bioavailability and bioactivity of select peptides. Our results suggest that O-glycosylation plays much wider roles in regulating and tuning the functions of peptide hormones than currently acknowledged, and considering O-glycans in studies, biomarker assays, and rational drug designs may be fruitful as recently demonstrated for ANP.

## Methods

**Ethical considerations.** Rat in vivo studies and extraction of porcine tissues were conducted with permission from the Danish Animal Experiments Inspectorate (2013-15-2934-00833), and in accordance with the guidelines of Danish legislation governing animal experimentation (1987) and the National Institutes of Health (publication number 85-23). The Regional Committee on Health Research Ethics approved the use of human tissue (KF 01287197), written informed consent was obtained from all study participants, and the study abides by the Declaration of Helsinki principles. Plasma and CSF samples were collected with written informed consent, and subsequently pooled and anonymized. The National Committee on Health Research Ethics has evaluated that the use of these samples for the glycoproteomics study did not need approval because of the anonymization of patient samples.

**Tissue extraction.** Porcine brain, cerebellum, ileum, and pancreas tissue from female LYD pigs of ~30 kg, as well as atrial appendage tissue from newborn piglets were extracted according to standard protocols. Prostate tissue specimens were collected from 25 patients undergoing radical prostatectomy for cancer (median age 65 years (range 54–72))<sup>70</sup>. Aliquots of frozen tissue were crushed using a CryoPrep tissue extractor (Covaris, Woburn, Massachusetts) and proteins were extracted in boiling water for 20 min, followed by homogenization with an Ultra-Turrax homogenizer (IKA, Staufen, Germany; Supplementary Data 2 and 3: LMWE H<sub>2</sub>O). After 30 min centrifugation at 13,000 × g, supernatants were collected. The ileum sample pellet was further extracted in 0.5 M acetic acid (AcOH) at room temperature (RT) for 30 min (Supplementary Data 2 and 3: LMWE AcOH) followed by centrifugation as above. For acidic ethanol extractions, ileum and pancreatic tissues were homogenized and extracted by rotating at 4 °C for 4 h in 0.18 M HCl/70% ethanol (Supplementary Data 2 and 3: LMWE EtOH/HCl). The samples were adjusted to pH 7 (NaOH), followed by centrifugation as above. A fraction of the brain, cerebellum, prostate gland, and ileum water extract underwent additional low-molecular-weight enrichment by either acetone precipitation (Supplementary Data 2 and 3: LMWE acetone) or acetone precipitation after acidification (Supplementary Data 2 and 3: LMWE acetone AcOH). Precipitation was performed either before or after adjusting to 0.5 M AcOH by addition of ice cold acetone (final v/v 67%), followed by incubation for 1 h at –20 °C, and subsequent centrifugation at 16,000 × g for 30 min (supernatants were collected). All supernatants were lyophilized and reconstituted in Milli-Q water followed by protein concentration determination by Pierce BCA protein assay kit (Pierce, Thermo Fischer Scientific), except the water extracts where lyophilization and reconstitution was omitted.

Rats were sacrificed and pancreas, heart, cerebellum, and brain tissues were dissected free and snap-frozen in liquid nitrogen. Frozen tissues were pulverized

using a mortar and pestle in liquid nitrogen. Proteins were extracted from ~100 mg tissue in 300  $\mu$ L 50 mM ammonium bicarbonate containing 1% RapiGest SF Surfactant (Waters, from here on: Rapigest), using an IKA Ultra Turbax blender at maximum speed for 20 s followed by 30 s of sonication using a Sonic Dismembrator (Fischer Scientific). Samples were diluted to 0.1% Rapigest in 50 mM ammonium bicarbonate and immediately subjected to enzymatic digestion and desialylation.

**Enzyme digestion and desialylation.** All extracted samples were adjusted to 50 mM ammonium bicarbonate, heated for 10 min at 80 °C, followed by reduction with 5 mM dithiothreitol (60 °C, 30 min) and alkylation with 10 mM iodoacetamide (30 min, RT, in the dark). Subsequently, samples were incubated with trypsin, chymotrypsin or GluC (Roche; 37 °C, overnight, 1  $\mu$ g enzyme per 100  $\mu$ g protein). The following day, the enzyme reaction was quenched and Rapigest was precipitated by acidifying with trifluoroacetic acid (TFA). The solution was cleared by centrifugation (10,000  $\times$  g, 10 min) and peptides were purified on C18 Sep-Pak columns (Waters), and dried down using a SpeedVac vacuum concentrator (Thermo Fischer Scientific). If not already desialylated, the dried peptides were resuspended in 1 mL 50 mM sodium acetate (pH 5.5) containing 0.1 U/mL neuraminidase (Sigma, N3001) followed by incubation at 37 °C for 1 h, purified by C18 Sep-Pak columns and dried down.

In the case of rat tissues, 200  $\mu$ g digest was labeled with TMTsixplex or TMT10plex (Thermo Fischer Scientific) according to manufacturer's instructions (Supplementary Data 2 and 3: Rapigest).

**Cell protein extraction.** STC-1 (CRL-3254) cells derived from mouse invasive small intestinal neuroendocrine carcinoma and N2a (CCL-131) cells from mouse neuroblastoma were obtained from ATCC and maintained in Dulbecco's modified Eagle's medium (DMEM) (Sigma) supplemented with 10% fetal bovine serum (Sigma) and 4 mM GlutaMAX (Gibco) at 37 °C and 5% CO<sub>2</sub>. Conditioned media cleared from dead cells and debris obtained from 2  $\times$  T175 flasks cultured for 48–72 h were dialyzed, neuraminidase (N3001, *Clostridium perfringens* neuraminidase type VI, Sigma) treated, and enriched for glycoprotein by capture on short (300  $\mu$ L beads contained in 1 mL syringes) PNA agarose columns (Vector labs). Glycoproteins were eluted by heating the lectin (2  $\times$  90 °C, 10 min) in 50 mM ammonium bicarbonate containing 0.05 % Rapigest. Total cell lysates were obtained by washing a monolayer of cells in ice cold phosphate-buffered saline (PBS), scraping off the cells and adding 2 mL 0.05–0.1% Rapigest to solubilize the cell pellet. The resulting suspension was sonicated and cleared by centrifugation followed by enzymatic digestion and desialylation as described above.

**Biofluid extraction.** Blood for O-glycoproteomic plasma analyses was collected and pooled from two healthy volunteers into EDTA-coated tubes (K2E K2EDTA Vacuette) followed by centrifugation at 5000  $\times$  g for 10 min and stored at –20 °C until use. CSF was obtained and pooled from 50 patients that underwent lumbar puncture. The samples were anonymized before use. Both crude and LMWF-enriched plasma and CSF were subjected to the O-glycoproteome workflow. For LMWF enrichment, 40 mL CSF or a volume of plasma containing ~60 mg of proteins (measured by Pierce BCA protein assay kit) was precipitated by adding two parts of 96% ethanol followed by incubation at RT for 30 min. Samples were centrifuged at 10,000  $\times$  g for 10 min, and the supernatant was lyophilized and resuspended in 0.05% Rapigest followed by desialylation in 50 mM sodium acetate buffer with 0.1 U/mL neuraminidase (Sigma, N3001). The LMWF-enriched samples were further enriched for O-glycopeptides on short PNA columns as done for the cell secretomes. In parallel, 5 mg of total protein (as determined by Pierce BCA protein assay kit) from non-LMWF-enriched biofluid samples was desialylated as described for the LMWF-enriched samples, before enzymatic degradation omitting the glycoprotein enrichment on short PNA columns.

**LWAC O-glycopeptide enrichment.** Agarose bound lectins PNA (binds galactosyl ( $\beta$ -1,3) N-acetylgalactosamine, T-antigen), Jacalin (binds galactosyl ( $\beta$ -1,3) N-acetylgalactosamine, T-antigen, or  $\alpha$ -N-acetylgalactosamine, Tn-antigen<sup>71</sup>), or VVA ( $\alpha$ -N-acetylgalactosamine, Tn-antigen) were obtained from Vector Laboratories. Dried samples were reconstituted in 2 mL lectin-binding buffer (PNA-binding buffer 10 mM HEPES (pH 7.4), 150 mM NaCl, 0.1 mM CaCl<sub>2</sub>, and 0.01 mM MnCl<sub>2</sub>; Jacalin-binding buffer 175 mM Tris (pH 7.5); VVA-binding buffer 20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM CaCl<sub>2</sub>/MgCl<sub>2</sub>/MnCl<sub>2</sub>/ZnCl<sub>2</sub>, and 1 M urea), filtered through a 0.45  $\mu$ m spin column and injected onto a pre-equilibrated 2.6 m long column packed with lectin-bound (PNA, Jacalin, or VVA) agarose beads at a constant flow rate of 0.1 mL/min. For VVA (N2A cells are naturally deficient in O-glycan elongation and produce glycoproteins with truncated Tn (HexNAc) glycans), the column was first washed for 3  $\times$  column volumes (CV) in 0.4 M glucose and glycopeptides were eluted with 2  $\times$  CV 0.2 M GalNAc and 1  $\times$  CV 0.4 M GalNAc. For PNA and Jacalin LWAC, the column was washed 2  $\times$  CV in lectin-binding buffer, and glycopeptides were eluted with 2  $\times$  CV 0.5 M galactose and 1  $\times$  CV 1 M galactose, respectively. A total of 5% of the elution fractions were analyzed by LC-MS/MS to check for sample complexity. If appropriate, the elution fractions were pooled and fractionated using either high pH-fractionation<sup>52</sup> or isoelectric focusing prior to nLC-MS/MS analysis.

**nLC-MS/MS analysis.** LC-MS/MS was performed on a system composed of an EASY-nLC 1000 (Thermo Fisher Scientific) interfaced via a nanoSpray Flex ion source to an LTQ-Orbitrap Velos pro hybrid spectrometer or Orbitrap Fusion Tribrid (Thermo Fisher Scientific), equipped for both higher energy collisional dissociation (HCD) and electron transfer dissociation (ETD) modes, enabling peptide sequence analysis with and without retention of glycan site-specific fragments, respectively. The nLC was operated in a one-column setup with an analytical column (20 cm length, 75  $\mu$ m inner diameter) packed with C18 reverse phase material (1.9- $\mu$ m particle size, ReproSil-Pur, Dr. Maisch, Ammerbuch Entringen, Germany). Each sample dissolved in 0.1% formic acid was injected onto the column and eluted in a gradient from 2% to 30% B in 105 min (2 h method) or 165 min (3 h method), from 30% to 100% B in 5 min and 100% B for 10 min at 200 nL min<sup>-1</sup> (solvent A, 100% H<sub>2</sub>O; solvent B, 100% acetonitrile; both containing 0.1% (v/v) formic acid). A data-dependent mass spectral acquisition routine, ETD triggering of subsequent HCD scan, was used for all runs. Briefly, a precursor MS1 scan ( $m/z$  350–1700) of intact peptides was acquired in the Orbitrap at a resolution setting of 30,000 (Velos Pro) or 120,000 (Fusion), followed by Orbitrap ETD-MS2 at a resolution setting of 15,000 (Velos Pro) or 60,000 (Fusion;  $m/z$  of 120–2000) of the five most abundant multiply charged precursors in the MS1 spectrum; this event was followed up by an HCD-MS2 fragmentation at a resolution setting of 7500 (Velos Pro) or 60,000 (Fusion;  $m/z$  of 120–2,000) for the same precursor ion. In cases where preliminary screening of fractions for glycopeptide enrichment was carried out prior to IEF, the ETD-MS2 step was omitted, and HCD-MS2 at a resolution setting of 7500 (Velos Pro) or 60,000 (Fusion;  $m/z$  120–2000) of the ten most abundant multiply charged precursors was acquired (top ten method). These HCD-MS2 spectra were simply screened for the appearance of the HexNAc oxonium fragment ions at  $m/z$  204.086.

**Data analysis.** The raw data were processed using Proteome Discoverer 1.4 software (Thermo Fisher Scientific) and searched against the human, porcine, mouse, or rat-specific Uniprot databases downloaded in January 2013. The Sequest HT search node was used for HCD and ETD data. In all cases, the precursor mass tolerance was set to 15 p.p.m. and fragment ion mass tolerance to 20 millimass units. Carbamidomethylation on cysteine residues was used as a fixed modification. Methionine oxidation, and HexNAc or HexHexNAc attachment to serine, threonine, or tyrosine were used as variable modifications. In the case of the VVA-enriched N2a cell line samples only HexNAc (Ser, Thr, and Tyr) and methionine oxidation were used as variable modifications. As an additional preprocessing procedure, all HCD spectra showing the presence of fragment ions at  $m/z$  204.08 were extracted into a single.mgf file, and the exact mass of 1 $\times$ , 2 $\times$ , 3 $\times$ , and 4 $\times$  HexNAc, Hex(1)HexNAc(1) units or the different combinations of the two modifications were subtracted from the corresponding precursor ion mass, generating a distinct file for each subtracted glycan mass<sup>38</sup>. In this process, a suffix of \_#xT and/or \_#xTn was added to the filename of the generated.mgf files, indicating the number of subtracted masses of Hex(1)HexNAc(1) and/or HexNAc units, respectively. In the case of VVA-enriched samples, the subtraction was only performed for HexNAc, resulting in a \_#xGalNAc suffix added to the.mgf filename. These preprocessed data files were submitted to a Sequest HT node under the conditions mentioned above, without considering HexNAc or Hex(1)HexNAc(1) modifications. Unassigned spectra were submitted to a second Sequest HT node using the same parameters as above with the exception of performing the search using semi-specific trypsin/chymotrypsin/GluC cleavage; all spectra were searched against a decoy database using a target false discovery rate of 1%. The final list was filtered against the human part of the NeuroPep database<sup>24</sup> to include only peptide hormones.

**Multiple sequence alignment.** All alignments were performed in ClustalO using the peptide sequences of *Homo sapiens*, *Sus scrofa*, *Mus musculus*, and *Rattus norvegicus*.

**Immunoblotting.** PreproNPY-plasmid (Human, Uniprot: P01303) was obtained from Genewiz ([www.genewiz.com](http://www.genewiz.com)) in the pUC57 framework and subcloned into pCDNA3-plasmid (Invitrogen, Thermo Fischer Scientific) with Bam/NotI restriction enzymes (New England Biolabs) followed by ligation with T4 DNA ligase (Thermo Fisher Scientific). STC-1 cells<sup>72</sup> stably transfected with preproNPY (maintained in 700  $\mu$ g/mL Zeocin (Invitrogen, Sigma)) were seeded out in six-well dishes and grown to 50% confluency. Cells were incubated in Opti-Mem media (Sigma) for 72 h to accumulate secreted proteins.

HEK293-6E (obtained through a license agreement with Dr. Yves Durocher, Bioprocédés Institute de recherche en biotechnologie, Montréal) were grown in suspension in serum-free F17 medium (Invitrogen, Thermo Fischer Scientific) supplemented with 0.1% Kolliphor P188 (Sigma) and 4 mM GlutaMax (Gibco, Sigma) at 37 °C and 5% CO<sub>2</sub> under constant agitation (120 r.p.m.). A total of 15 mL of HEK293-6E cells (1  $\times$  10<sup>6</sup> cells/mL) were transfected with preproNPY by adding combined 15  $\mu$ g DNA and 90  $\mu$ g of polyethylenimine-25K (Polysciences, Warrenton, PA, USA), co-incubated at RT for 10 min in 500  $\mu$ L Opti-Mem. HEK293-6E cells were cultured for 5 days before harvest of the conditioned media.

For enzymatic treatment, the conditioned medium was adjusted to 20 mM Tris (pH 6.8). STC-1 samples were treated with 0.05 mU/ $\mu$ L neuraminidase (Sigma, N3001) overnight (37 °C), and HEK-6E samples were treated with 0.05 mU/ $\mu$ L neuraminidase (Sigma, N3001) and 4 U/ $\mu$ L OPERATOR for 6 h (37 °C).



Subsequently, 200  $\mu$ L of each STC-1 sample was concentrated through trichloroacetic acid precipitation (final v/v 10%) followed by centrifugation (16,000  $\times$  g, 10 min), and resuspension of the pellet in 15  $\mu$ L 67 mM Tris (pH 8), 1  $\times$  NuPage<sup>®</sup> LDS-buffer (Thermo Fischer Scientific) and 1  $\times$  Bolt Sample Reducing Agent (Thermo Fischer Scientific) to allow for immunoblotting of all 200  $\mu$ L/lane. A total of 10  $\mu$ L of each HEK-6E sample was stopped by addition of 3.5  $\mu$ L 4  $\times$  LDS buffer and 1.5  $\mu$ L 10  $\times$  Bolt Sample Reducing Agent to a final volume of 15  $\mu$ L and blotted without upconcentration.

Samples were boiled 10 min at 80  $^{\circ}$ C, prior to separation on 4–12% Bis-tris gels (NuPage, Thermo Fischer Scientific) and blotted onto nitrocellulose membrane for 11 min at 20 V, using the semi-dry Iblot2 system (Thermo Fischer Scientific). Antibody staining was performed according to the modified protocol by Okita et al.<sup>73</sup>. Membranes were blocked 5 min in 1% milk, 0.1% bovine serum albumin (Sigma; w/v) in Tris-buffered saline with 0.05% Tween-20 (TBS-T) followed by 3 min wash in PBS with 0.05% Tween-20 (PBS-T). Blots were fixed in 0.2% glutaraldehyde (Sigma) in PBS-T for 15 min at RT. After three brief washes in PBS-T, antigen retrieval was performed by microwave treatment (600 W, 10 min) in citrate retrieval buffer (10 mM citric acid, 1 mM EDTA, 0.05% Tween-20, pH 6.0), followed by incubation in quenching buffer for 10 min at RT (200 mM glycine in TBS-T, pH 7.4). Blots were incubated overnight with anti-NPY antibody (Cell Signaling Technologies D7Y5A, Rabbit mAb #11976) diluted 1:1000 in the blocking solution at 4  $^{\circ}$ C. On the next day, the blots were washed in TBS-T followed by 1 h incubation with secondary HRP-conjugated goat antirabbit Ig's (DAKO, P0448, 0.25  $\mu$ g/mL). The membranes were washed in TBS-T and developed using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Fischer Scientific). Pictures were taken with a LAS-3000 instrument (GE healthcare).

**Purified glycopeptides.** Peptide hormones were obtained from chemical synthesis (Synpeptides) in both non-glycosylated and T-glycosylated forms, except the VIP and SCT peptides that were only obtained in non-glycosylated forms. NPY and PYY were also obtained in their Tn forms (Synpeptides). To obtain Tn-glycosylated variants of VIP, GCG, GAL, SCT, and GLP-1, non-glycosylated peptides were chemoenzymatically glycosylated with GalNAc-T1<sup>25</sup> (25 mM cacodylic acid sodium, pH 7.4, 10 mM MnCl<sub>2</sub>, 0.25% Triton X-100, 2 mM UDP-GalNAc (Sigma), 40  $\mu$ g/mL enzyme, and 1 mg/mL peptide, 37  $^{\circ}$ C overnight) and C18 high-performance liquid chromatography (HPLC) purified. Correct site incorporation was confirmed by LC-MS. The Tn-glycosylated peptides obtained by either chemical or chemoenzymatic synthesis were chemoenzymatically elongated to the T-variant using the T-synthase<sup>74</sup> (25 mM cacodylic acid sodium, pH 7.4, 10 mM MnCl<sub>2</sub>, 0.25% Triton X-100, 2 mM UDP-Gal (Sigma), 4 pmol/h/ $\mu$ L enzyme, and 1 mg/mL peptide, 37  $^{\circ}$ C overnight) and HPLC purified. A fraction of all T peptides was chemoenzymatically elongated to the ST form using purified ST3Gal1 (50 mM MES pH 6.5, 2 mM CMP-NeuAc (Sigma), 40  $\mu$ g/mL enzyme, 1 mg/mL peptide, 37  $^{\circ}$ C overnight)<sup>75</sup>. The glycoproteins from each reaction were separated from non-glycosylated product on C18 HPLC (Thermo, Ultimate<sup>™</sup> 3000, Phenomenex, Jupiter, 5  $\mu$ m, 300  $\text{\AA}$ , 250 mm) under isocratic conditions (between 20% and 40% acetonitrile dependent on peptide), lyophilized and resuspended in Milli-Q water. Glycopeptide product quantification was performed using HPLC (UV 210 nm) by comparison with a standard curve of weighed non-glycosylated peptide as standard. GCG<sup>Thr7-T</sup> was deglycosylated using O-glycanase (Merck, #324716; 50 mM phosphate buffer pH 6.0, 0.13 mU/mL enzyme, 1 mg/mL GCG<sup>Thr7-T</sup>, 37  $^{\circ}$ C for 2 h) and HPLC purified as described above. Deglycosylated GCG precipitated at assay conditions, and was resolubilized by addition of 0.1% TFA prior to HPLC purification.

**Transfection and tissue cultures for receptor activation.** COS-7 (CRL-1651, ATCC) cells were grown in DMEM 1885 supplemented with 10% fetal bovine serum (Sigma), 2 mM glutamine, 180 units/ml penicillin, and 45 g/ml streptomycin at 10% CO<sub>2</sub> and 37  $^{\circ}$ C. One day after seeding in culture flasks (seeding density of 6 million cells per T175 flask), the cells were transiently transfected with 40  $\mu$ g receptor DNA for cAMP determination and a combination of 20  $\mu$ g receptor DNA and 30  $\mu$ g of pcDNA 1.1 plasmid encoding a chimeric G protein Ga $\Delta$ 6q4myr<sup>76</sup> for IP3 determination, using the calcium phosphate precipitation method<sup>77</sup>.

**IP3 assay.** One day after transfection, COS-7 cells transiently expressing either NPY1R, 2R, 4R, or 5R (35,000 cells/well) were incubated with myo-[<sup>3</sup>H]inositol (5  $\mu$ L/ml, 2  $\mu$ Ci/ml) in 0.1 ml of medium overnight in a 96-well plate. The following day, cells were washed twice in PBS and incubated in 0.1 ml of Hanks' balanced salt solution (Invitrogen, Thermo Fischer Scientific) supplemented with 10 mM LiCl at 37  $^{\circ}$ C in the presence of various concentrations of NPY or PYY, and their respective glycoforms for 90 min (in duplicate determinations). Assay medium was then removed, and cells were extracted by the addition of 50  $\mu$ L of 10 mM formic acid to each well, followed by incubation on ice for 30–60 min. The [<sup>3</sup>H]inositol phosphates in the formic acid cell lysates were thereafter quantified by adding yttrium silicate-poly-D-Lys-coated SPA beads<sup>78</sup>. Briefly, 35  $\mu$ L of cell extracts were mixed with 80  $\mu$ L of SPA bead suspension in H<sub>2</sub>O (12.5  $\mu$ g/ $\mu$ L) in a white 96-well plate. Plates were sealed and shaken on table shaker for at least 30 min. SPA beads were allowed to settle and react with the extract for at least 8 h before radioactivity was determined using a Packard Top Count NXT scintillation counter (PerkinElmer Life Sciences).

**cAMP assay.** One day after transfection, the COS-7 cells were seeded at 35,000 cells/well in 96-well plates. 24 h later, cells were washed twice with HEPES-buffered saline (in mmol/L: 20 HEPES, 150 NaCl, 0.75 NaH<sub>2</sub>PO<sub>4</sub>, pH 7.4) and incubated 30 min at 37  $^{\circ}$ C in HEPES-buffered saline containing 1 mmol/L iso-butylmethylxanthine (IBMX) phosphodiesterase inhibitor (Sigma-Aldrich, Brøndby, Denmark). Peptide hormones, either non-glycosylated or glycosylated on Thr7, were then added (in duplicate determinations) followed by 30 min incubation at 37  $^{\circ}$ C. Subsequently, the medium was removed and the cells were treated with the enzyme fragment complementation-based cAMP assay according to manufacturer's instructions (HitHunter cAMP XS+assay, DiscoverX, Fremont, CA). The cAMP content was measured as luminescence using an EnVision 2104 Multititle Platerreader (PerkinElmer) with a cAMP standard curve for validation.

**Enzyme assay.** Recombinant protease assays were performed in a MALDI-TOF timecourse assay. A total of 15  $\mu$ M peptide hormone underwent enzymatic treatment in vitro at 37  $^{\circ}$ C in either 50 mM Tris, 0.05% Brij-35, pH 9.0 for NEP assays, or 50 mM Tris, pH 8.0 for DPP-IV assays. The amount of enzyme used was optimized to fully digest the non-glycosylated peptide hormone of interest within 1 h of incubation. The final concentrations used in the assays are as follows; NEP (R&D Systems): 150  $\mu$ g/ $\mu$ L enzyme for VIP, GAL, and SCT assays. A total of 10 ng/ $\mu$ L for GLP-1 assays, and 20 ng/ $\mu$ L for PYY assays; DPP-IV (R&D Systems): 4 ng/ $\mu$ L for VIP assays, 5 ng/ $\mu$ L for GLP-1 assays, and 10 ng/ $\mu$ L for PYY assays. Product development was evaluated by quenching 1  $\mu$ L reaction in 40  $\mu$ L 0.1% TFA after 15 min, 30 min, 60 min, 120 min, and 24 h of reaction time, followed by MALDI-TOF analysis. A 1:1 mixture of GLP-1<sup>Thr5</sup> and GLP-1<sup>Thr7</sup> glycosylated variants was used in this assay due to limited amounts of material.

MALDI-TOF-MS was performed in linear positive mode on a Bruker Autoflex instrument (Bruker Daltonik GmbH, Bremen, Germany) by mixing the quenched aliquots with a saturated solution of  $\alpha$ -Cyano-4-hydroxycinnamic acid in ACN/H<sub>2</sub>O/TFA (70:30:0.1) at a 1:1 ratio on a target steel plate.

**CD spectroscopy.** CD spectra of 20  $\mu$ M NPY or NPY glycovariants were obtained in 15 mM sodium phosphate buffer (pH 7) at 25  $^{\circ}$ C between 200 and 260 nm at a scan rate of 50 nm/min. Each spectrum was obtained as an average of 15 accumulations corrected by subtracting the measurements in 15 mM phosphate buffer alone. Measurements were performed in a Jasco-810 spectropolarimeter equipped with Peltier temperature control, using a quartz cuvette of 0.1 cm path length. Alpha-helical content was calculated by the mean residue molar ellipticity at 222 nm as  $[-[\theta]_{222} + 3000]/(36,000 + 3000) \times 100$ <sup>79</sup>.

**Half-life of native and glycosylated GLP-1 isoforms in rats.** Male Wistar rats ( $\approx$ 250 g) were obtained from Janvier (Saint Berthevin Cedex, France) and housed in pairs under standard conditions with ad libitum access to chow and water. Rats were acclimatized for at least 1 week before the study. Studies were carried out on two occasions, matching on each occasion the number of rats in each group, while at the same time avoiding rats from the same cage receiving the same test solution. Rat body weight did not differ between treatment groups ( $343 \pm 14.36$  g vs.  $333.9 \pm 15.74$  g vs.  $353.3 \pm 12.57$  g,  $p > 0.359$ ). On the day of study, non-fasted rats were anesthetized by a subcutaneous injection with Hypnorm/Midazolam (0.3 ml per 100 g body weight) and placed on a heated operating table (37  $^{\circ}$ C). The abdominal cavity was opened by a midline incision and a plastic catheter was inserted in the portal vein. Blood was collected ( $\sim$ 300  $\mu$ L per timepoint) through the catheter into ice cold EDTA-coated tubes that were immediately placed on ice. Samples were collected at timepoint -5, 0, 1, 2, 5, 10, 15, 25, 45, and 60 min. In between samples,  $\sim$ 200  $\mu$ L PBS were injected to flush the catheter and replace lost fluid. Total blood volume withdrawal was  $\sim$ 3 ml, corresponding to  $\sim$ 15% of the total blood volume (estimated by: blood volume (ml) =  $0.06 \times$  body weight (g) +  $0.77$ )<sup>80</sup>. Immediately after the 0 min sample collection, one of three test compounds was injected (20 nmol per kg body weight) into the inferior vena cava through a 26 Gauge needle. Test compounds consisted of synthetic GLP-1 7–36 amide and two glycosylated GLP-1<sup>Thr7-T</sup> and GLP-1<sup>Thr7-ST</sup> isoforms, prepared in PBS + 1% Human Serum Albumin (Millipore) and injection volume was  $\sim$ 300  $\mu$ L. Blood was centrifuged (1650  $\times$  g, 4  $^{\circ}$ C, 15 min) within 30 min after collection, and plasma was transferred into clean Eppendorf tubes and stored at -20  $^{\circ}$ C until hormone analysis. At end of study, rats were sacrificed by diaphragm perforation.

**Biochemical measurements.** Plasma concentrations of total GLP-1 were measured using an in-house assay (code: 89390), employing an antibody diluted 1:200,000 specifically targeting the amidated C-terminus of the molecule, thus measuring both intact GLP-1 (7–36 amide), the primary metabolite (9–36 amide) and other potential N- or mid-terminally truncated isoforms<sup>81</sup>. Non-extracted samples were diluted 500 $\times$  in assay buffer to stay within the sensitive range of the assay, and to eliminate potential cross-reaction between the antibodies and non-GLP-1 related plasma molecules (matrix effect).

**Statistical analysis.** All parameters and errors were determined with the GraphPad Prism version 8 software, except half-life determinations that were performed with the R Studio software.

Within each independent receptor activation assay, a nonlinear regression (three parameter logistic fit with Hill coefficient = 1) was fitted to the raw data of the non-glycosylated agonist, and top (100% receptor activation) and bottom (0% receptor activation) parameters were used for normalization of all data within that assay. LogEC50's of the normalized values were determined by nonlinear regression (constraining the Hill coefficient = 1 and the top ( $E_{\max}$ ) to 100% receptor activation). For  $E_{\max}$  determinations, the same analysis was repeated without constraint on the top parameter.

Half-life was determined individually for each GLP-1 variant using the linear model command (lm, Concentration ~ Timepoint) on ln-transformed data. Only data points represented by  $n > 4$  were included in the analysis. To test the difference in rate constant ( $k$ ), a multiple linear model (lm) was fitted to the pooled data with glycoform set as interaction (lm, Concentration ~ Timepoint \* Glycoform). The analysis was repeated while removing the non-glycosylated peptide data to test for difference between the GLP-1 T and ST forms (\*\* $p < 0.001$ ; \* $p < 0.01$ ; \* $p < 0.05$ ).

When comparing multiple groups, one-way ANOVA was performed with Tukey's post hoc test for testing difference between multiple groups (ns, not significant; \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$  (two tailed)). Comparison of exactly two groups was performed with two-tailed Student's  $t$ -test (ns, not significant; \*\*\* $p < 0.001$ ; \*\* $p < 0.01$ ; \* $p < 0.05$  (two tailed)).

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The MS proteomics data underlying Fig. 1, Table 1, Figs. 2 and 3, and Supplementary Figs. 1–3 have been deposited to the ProteomeXchange Consortium via the MassIVE partner repository ([massive.ucsd.edu](https://massive.ucsd.edu)) with the dataset identifier PXD018560. Data are also available at the MassIVE repository with identifier MSV000085289. Source data are provided with this paper. All other data are available in the main text, the Supplementary Materials, or from the corresponding author on reasonable request.

Received: 10 October 2019; Accepted: 2 July 2020;

Published online: 20 August 2020

## References

- Clemmensen, C. et al. Emerging hormonal-based combination pharmacotherapies for the treatment of metabolic diseases. *Nat. Rev. Endocrinol.* **15**, 90–104 (2019).
- Fosgerau, K. & Hoffmann, T. Peptide therapeutics: current status and future directions. *Drug Discov. Today* **20**, 122–128 (2015).
- Erak, M., Bellmann-Sickert, K., Els-Heindl, S. & Beck-Sickinger, A. G. Peptide chemistry toolbox—transforming natural peptides into peptide therapeutics. *Bioorg. Med. Chem.* **26**, 2759–2765 (2018).
- Wei, S., Feng, Y., Kalinina, E. & Fricker, L. D. Neuropeptide-processing carboxypeptidases. *Life Sci.* **73**, 655–662 (2003).
- Seidah, N. G. What lies ahead for the proprotein convertases? *Ann. N. Y. Acad. Sci.* **1220**, 149–161 (2011).
- Milgram, S. L., Mains, R. E. & Eipper, B. A. COOH-terminal signals mediate the trafficking of a peptide processing enzyme in endocrine cells. *J. Cell Biol.* **121**, 23–36 (1993).
- Bundgaard, J. R., Vuust, J. & Rehfeld, J. F. Tyrosine O-sulfation promotes proteolytic processing of progastrin. *EMBO J.* **14**, 3073–3079 (1995).
- Drazic, A., Myklebust, L. M., Ree, R. & Arnesen, T. The world of protein acetylation. *Biochim. Biophys. Acta* **1864**, 1372–1401 (2016).
- Secher, A. et al. Analytic framework for peptidomics applied to large-scale neuropeptide identification. *Nat. Commun.* **7**, 11436 (2016).
- Hansen, L. H. et al. Discovery of O-glycans on atrial natriuretic peptide (ANP) that affect both its proteolytic degradation and potency at its cognate receptor. *J. Biol. Chem.* **294**, 12567–12578 (2019).
- Mentlein, R. Dipeptidyl-peptidase IV (CD26)—role in the inactivation of regulatory peptides. *Regul. Pept.* **85**, 9–24 (1999).
- Potter, L. R. Natriuretic peptide metabolism, clearance and degradation. *Febs J.* **278**, 1808–1817 (2011).
- Deacon, C. F., Nauck, M. A., Meier, J., Hucking, K. & Holst, J. J. Degradation of endogenous and exogenous gastric inhibitory polypeptide in healthy and in type 2 diabetic subjects as revealed using a new assay for the intact peptide. *J. Clin. Endocrinol. Metab.* **85**, 3575–3581 (2000).
- Lietz, C. B. et al. Phosphopeptidomics reveals differential phosphorylation states and novel SxE phosphosite motifs of neuropeptides in dense core secretory vesicles. *J. Am. Soc. Mass Spectrom.* **29**, 935–947 (2018).
- Stentoft, C. et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* **32**, 1478–1488 (2013).
- Schjoldager, K. T. et al. Deconstruction of O-glycosylation–GalNAc-T isoforms direct distinct subsets of the O-glycoproteome. *EMBO Rep.* **16**, 1713–1722 (2015).
- Narimatsu, Y. et al. Exploring regulation of protein O-glycosylation in isogenic human HEK293 cells by differential O-glycoproteomics. *Mol. Cell Proteom.* **18**, 1396–1409 (2019).
- Yu, Q. et al. Targeted mass spectrometry approach enabled discovery of O-glycosylated insulin and related signaling peptides in mouse and human pancreatic islets. *Anal. Chem.* **89**, 9184–9191 (2017).
- Cao, F. et al. Detection of biosynthetic precursors, discovery of glycosylated forms, and homeostasis of calcitonin in human cancer cells. *Anal. Chem.* **89**, 6992–6999 (2017).
- Schjoldager, K. T. et al. A systematic study of site-specific GalNAc-type O-glycosylation modulating proprotein convertase processing. *J. Biol. Chem.* **286**, 40122–40132 (2011).
- Goth, C. K. et al. Proteolysis: a major role for regulated site-specific O-glycosylation. *Trends Biochem. Sci.* **43**, 269–284 (2018).
- Goth, C. K. et al. A systematic study of modulation of ADAM-mediated ectodomain shedding by site-specific O-glycosylation. *Proc. Natl Acad. Sci. USA* **112**, 14623–14628 (2015).
- Khetarpal, S. A. et al. Loss of function of GALNT2 lowers high-density lipoproteins in humans, nonhuman primates, and rodents. *Cell Metab.* **24**, 234–245 (2016).
- Wang, Y. et al. NeuroPep: a comprehensive resource of neuropeptides. *Database (Oxford)* **2015**, bav038 (2015).
- Kong, Y. et al. Probing polypeptide GalNAc-transferase isoform substrate specificities by in vitro analysis. *Glycobiology* **25**, 55–65 (2015).
- Portoghese, P. S. Bivalent ligands and the message-address concept in the design of selective opioid receptor antagonists. *Trends Pharm. Sci.* **10**, 230–235 (1989).
- Birch, N. P., Estivariz, F. E., Bennett, H. P. & Loh, Y. P. Differential glycosylation of N-POMC1-77 regulates the production of gamma 3-MSH by purified pro-opiomelanocortin converting enzyme. A possible mechanism for tissue-specific processing. *FEBS Lett.* **290**, 191–194 (1991).
- Hook, V. Y., Schiller, M. R. & Azaryan, A. V. The processing proteases prohormone thiol protease, PC1/3 and PC2, and 70-kDa aspartic proteinase show preferences among proenkephalin, proneurotensin, and proopioidmelanocortin substrates. *Arch. Biochem. Biophys.* **328**, 107–114 (1996).
- Brakch, N., Galanopoulou, A. S., Patel, Y. C., Boileau, G. & Seidah, N. G. Comparative proteolytic processing of rat prosomatostatin by the convertases PC1, PC2, furin, PACE4 and PC5 in constitutive and regulated secretory pathways. *FEBS Lett.* **362**, 143–146 (1995).
- Knudsen, S. M., Tams, J. W. & Fahrenkrug, J. Functional roles of conserved transmembrane prolines in the human VPAC(1) receptor. *FEBS Lett.* **503**, 126–130 (2001).
- Torang, S., Veefald, S., Rosenkilde, M. M., Hartmann, B. & Holst, J. J. The anorexic hormone Peptide YY3-36 is rapidly metabolized to inactive Peptide YY3-34 in vivo. *Physiol. Rep.* **3**, e12455 (2015).
- Holst, J. J., Albrechtsen, N. J. W., Gabe, M. B. N. & Rosenkilde, M. M. Oxyntomodulin: actions and role in diabetes. *Peptides* **100**, 48–53 (2018).
- Svendsen, B. et al. Insulin secretion depends on intra-islet glucagon signaling. *Cell Rep.* **25**, e1122 (2018).
- Nygaard, R., Nielbo, S., Schwartz, T. W. & Poulsen, F. M. The PP-fold solution structure of human polypeptide YY and human PYY3-36 as determined by NMR. *Biochemistry* **45**, 8350–8357 (2006).
- Merten, N. et al. Receptor subtype-specific docking of Asp6.59 with C-terminal arginine residues in Y receptor ligands. *J. Biol. Chem.* **282**, 7543–7551 (2007).
- de Graaf, C. et al. Extending the structural view of class B GPCRs. *Trends Biochem. Sci.* **42**, 946–960 (2017).
- Yang, Z. et al. Structural basis of ligand binding modes at the neuropeptide Y Y1 receptor. *Nature* **556**, 520–524 (2018).
- Lavery, S. B. et al. Advances in mass spectrometry driven O-glycoproteomics. *Biochim. Biophys. Acta*, **1850**, 33–42 (2014).
- Yang, S. et al. Deciphering protein O-glycosylation: solid-phase chemoenzymatic cleavage and enrichment. *Anal. Chem.* **90**, 8261–8269 (2018).
- King, S. L. et al. Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. *Blood Adv.* **1**, 429–442 (2017).
- Bader, R. & Zerbe, O. Are hormones from the neuropeptide Y family recognized by their receptors from the membrane-bound state? *Chembiochem* **6**, 1520–1534 (2005).
- Sargent, D. F. & Schwyzler, R. Membrane lipid phase as catalyst for peptide-receptor interactions. *Proc. Natl Acad. Sci. USA* **83**, 5774–5778 (1986).
- Nordmann, A., Blommers, M. J., Fretz, H., Arvinte, T. & Drake, A. F. Aspects of the molecular structure and dynamics of neuropeptide Y. *FEBS J.* **261**, 216–226 (1999).



44. Gerken, T. A., Butenhof, K. J. & Shogren, R. Effects of glycosylation on the conformation and dynamics of O-linked glycoproteins: carbon-13 NMR studies of ovine submaxillary mucin. *Biochemistry* **28**, 5536–5543 (1989).
45. Campbell, D. J. Long-term neprilysin inhibition—implications for ARNs. *Nat. Rev. Cardiol.* **14**, 171–186 (2017).
46. Lambeir, A. M., Durinx, C., Scharpe, S. & De Meester, I. Dipeptidyl-peptidase IV from bench to bedside: an update on structural properties, functions, and clinical aspects of the enzyme DPP IV. *Crit. Rev. Clin. Lab. Sci.* **40**, 209–294 (2003).
47. Parkes, D. et al. Pharmacokinetic actions of exendin-4 in the rat: Comparison with glucagon-like peptide-1. *Drug Dev. Res.* **53**, 260–267 (2001).
48. Bennett, E. P. et al. Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* **22**, 736–756 (2012).
49. Boskovski, M. T. et al. The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality. *Nature* **504**, 456–459 (2013).
50. Topaz, O. et al. Mutations in GALNT3, encoding a protein involved in O-linked glycosylation, cause familial tumoral calcinosis. *Nat. Genet.* **36**, 579–581 (2004).
51. Zilmer, M. et al. Novel congenital disorder of O-linked glycosylation caused by GALNT2 loss of function. *Brain* **143**, 1114–1126 (2020).
52. Hintze, J. et al. Probing the contribution of individual polypeptide GalNAc-transferase isoforms to the O-glycoproteome by inducible expression in isogenic cell lines. *J. Biol. Chem.* **293**, 19064–19077 (2018).
53. Luther, K. B., Schindelin, H. & Haltiwanger, R. S. Structural and mechanistic insights into lunatic fringe from a kinetic analysis of enzyme mutants. *J. Biol. Chem.* **284**, 3294–3305 (2009).
54. Wang, S. et al. Site-specific O-glycosylation of members of the low-density lipoprotein receptor superfamily enhances ligand interactions. *J. Biol. Chem.* **293**, 7408–7422 (2018).
55. Pedersen, N. B. et al. Low density lipoprotein receptor class a repeats are o-glycosylated in linker regions. *J. Biol. Chem.* **289**, 17312–17324 (2014).
56. Tian, E. et al. Galnt11 regulates kidney function by glycosylating the endocytosis receptor megalin to modulate ligand binding. *Proc. Natl Acad. Sci. USA* **116**, 25196–25202 (2019).
57. Taylor, M. E. & Drickamer, K. Mammalian sugar-binding receptors: known functions and unexplored roles. *Febs J.* **286**, 1800–1814 (2019).
58. Pedragosa-Badia, X., Stichel, J. & Beck-Sickinger, A. G. Neuropeptide Y receptors: how to get subtype selectivity. *Front. Endocrinol. (Lausanne)* **4**, 5 (2013).
59. Tagliabracci, V. S. et al. Dynamic regulation of FGF23 by Fam20C phosphorylation, GalNAc-T3 glycosylation, and furin proteolysis. *Proc. Natl Acad. Sci. USA* **111**, 5520–5525 (2014).
60. Yang, W. H. et al. An intrinsic mechanism of secreted protein aging and turnover. *Proc. Natl Acad. Sci. USA* **112**, 13657–13662 (2015).
61. Steirer, L. M., Park, E. I., Townsend, R. R. & Baenziger, J. U. The asialoglycoprotein receptor regulates levels of plasma glycoproteins terminating with sialic acid alpha2,6-galactose. *J. Biol. Chem.* **284**, 3777–3783 (2009).
62. Bilsky, E. J. et al. Enkephalin glycopeptide analogues produce analgesia with reduced dependence liability. *J. Med. Chem.* **43**, 2586–2590 (2000).
63. Yi, J., Warunek, D. & Craft, D. Degradation and stabilization of peptide hormones in human blood specimens. *PLoS ONE* **10**, e0134427 (2015).
64. Egleton, R. D. et al. Improved blood-brain barrier penetration and enhanced analgesia of an opioid peptide by glycosylation. *J. Pharm. Exp. Ther.* **299**, 967–972 (2001).
65. Meurer, J. A., Colca, J. R., Burton, P. S. & Elhammer, A. P. Properties of native and in vitro glycosylated forms of the glucagon-like peptide-1 receptor antagonist exendin(9-39). *Metab. Clin. Exp.* **48**, 716–724 (1999).
66. Dangoor, D. et al. Novel glycosylated VIP analogs: synthesis, biological activity, and metabolic stability. *J. Pept. Sci.* **14**, 321–328 (2008).
67. Kihlberg, J. et al. Glycosylated peptide hormones: pharmacological properties and conformational studies of analogues of [1-desamino,8-D-arginine] vasopressin. *J. Med. Chem.* **38**, 161–169 (1995).
68. Rosjo, H. et al. Influence of glycosylation on diagnostic and prognostic accuracy of N-terminal pro-B-type natriuretic peptide in acute dyspnea: data from the Akershus Cardiac Examination 2 Study. *Clin. Chem.* **61**, 1087–1097 (2015).
69. Saenger, A. K. et al. Specificity of B-type natriuretic peptide assays: cross-reactivity with different BNP, NT-proBNP, and proBNP peptides. *Clin. Chem.* **63**, 351–358 (2017).
70. Nielsen, S. J., Iversen, P., Rehfeld, J. F., Jensen, H. L. & Goetze, J. P. C-type natriuretic peptide in prostate cancer. *APMIS* **117**, 60–67 (2009).
71. Tachibana, K. et al. Elucidation of binding specificity of Jacalin toward O-glycosylated peptides: quantitative analysis by frontal affinity chromatography. *Glycobiology* **16**, 46–53 (2006).
72. Kuhre, R. E. et al. Peptide production and secretion in GLUTag, NCI-H716, and STC-1 cells: a comparison to native L-cells. *J. Mol. Endocrinol.* **56**, 201–211 (2016).
73. Okita, N. et al. Modified Western blotting for insulin and other diabetes-associated peptide hormones. *Sci. Rep.* **7**, 6949 (2017).
74. Ju, T., Brewer, K., D'Souza, A., Cummings, R. D. & Canfield, W. M. Cloning and expression of human core 1 beta1,3-galactosyltransferase. *J. Biol. Chem.* **277**, 178–186 (2002).
75. George, S. K. et al. Chemoenzymatic synthesis of sialylated glycopeptides derived from mucins and T-cell stimulating peptides. *J. Am. Chem. Soc.* **123**, 11117–11125 (2001).
76. Heydorn, A. et al. Identification of a novel site within G protein alpha subunits important for specificity of receptor-G protein interaction. *Mol. Pharm.* **66**, 250–259 (2004).
77. Kissow, H. et al. Glucagon-like peptide-1 (GLP-1) receptor agonism or DPP-4 inhibition does not accelerate neoplasia in carcinogen treated mice. *Regul. Pept.* **179**, 91–100 (2012).
78. Thiele, S. et al. Allosteric and orthosteric sites in CC chemokine receptor (CCR5), a chimeric receptor approach. *J. Biol. Chem.* **286**, 37543–37554 (2011).
79. Horn, M. & Neundorff, I. Design of a novel cell-permeable chimeric peptide to promote wound healing. *Sci. Rep.* **8**, 16279 (2018).
80. Lee, H. B. & Blaufox, M. D. Blood volume in the rat. *J. Nucl. Med.* **26**, 72–76 (1985).
81. Orskov, C., Rabenhøj, L., Wettergren, A., Kofod, H. & Holst, J. J. Tissue and plasma concentrations of amidated and glycine-extended glucagon-like peptide I in humans. *Diabetes* **43**, 535–539 (1994).

## Acknowledgements

This work was supported by the Novo Nordisk Foundation, the Lundbeck Foundation, Danish National Research Foundation Grant DNRF107.

## Author contributions

T.D.M. designed, planned, performed experiments, and wrote the manuscript. L.H.H. and R.E.K. designed and performed experiments. Z.Y. and S.Y.V. planned and performed experiments. S.J., D.B.A., H.J.J., and T.J. performed experiments. J.H., J.P.G., C.M., M.M.R., J.J.H., and C.G. designed experiments. K.T.S. designed, planned, and wrote the manuscript.

## Competing interests

The University of Copenhagen has filed patent applications EP2018/083325 partly on the basis of this work. T.D.M., L.H.H., S.Y.V., J.P.G., C.G., and K.T.S. are named inventors on this application. The other authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-020-17473-1>.

**Correspondence** and requests for materials should be addressed to K.T.S.

**Peer review information:** *Nature Communications* thanks Betty Eipper, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020